

TOWARDS ENSURING SCALABILITY,  
INTEROPERABILITY AND EFFICIENT ACCESS  
CONTROL IN A TRIPLE-DOMAIN GRID-BASED  
ENVIRONMENT

by

Nureni Ayofe Azeez

A thesis submitted in fulfillment of the requirements for



the degree of

Doctor of Philosophy in Computer Science

UNIVERSITY of the  
WESTERN CAPE

University of the Western Cape

November 2012

UNIVERSITY OF THE WESTERN CAPE

TOWARDS ENSURING SCALABILITY,  
INTEROPERABILITY AND EFFICIENT ACCESS  
CONTROL IN A TRIPLE-DOMAIN GRID-BASED  
ENVIRONMENT

by

Nureni Ayofe, Azeez

Supervisor:

Professor Tiko Iyamu

Co-Supervisor:

Professor Isabella M. Venter

**Abstract**

The high rate of grid computing adoption, both in academe and industry, has posed challenges regarding efficient access control, interoperability and scalability. Although several methods have been proposed to address these grid computing challenges, none has proven to be completely efficient and dependable. To tackle these challenges, a novel access control architecture framework, a triple-domain grid-based environment, modelled on role based access control, was developed. The architecture's framework assumes three domains, each domain with an independent Local Security Monitoring Unit and a Central Security Monitoring Unit that monitors security for the entire grid.

The architecture was evaluated and implemented using the G3S, grid security services simulator, meta-query language as "cross-domain" queries and Java Runtime Environment 1.7.0.5 for implementing the workflows that define the model's task. The simulation results show that the developed architecture is reliable and efficient if measured against the observed parameters and entities. This proposed framework for access control also proved to be interoperable and scalable within the parameters tested.

**Categories and Subject Descriptors**

D.4.6 [Security and Protection]: access controls, authentication, information flow control, interoperability, unauthorized access, scalability

**General Terms**

Security

**Keywords:** grid computing, computational grid, sensitive information, authorization, authentication, interoperability, scalability, security model, triple domain grid based environment.



## TABLE OF CONTENTS

<b>Table of Contents</b>	<b>iii</b>
<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>viii</b>
<b>List of Publications</b>	<b>ix</b>
<b>Acknowledgments</b>	<b>xi</b>
<b>Declaration</b>	<b>xv</b>
<b>Glossary</b>	<b>xvi</b>
<b>CHAPTER 1</b>	<b>1</b>
<i>Statement and analysis of the problem</i>	<b>1</b>
Sketching the background	1
Definition of a grid	1
Why secure the grid?	2
Security challenges to grid computing	3
Scalability	3
Interoperability	4
Efficient access control:	4
Research framework	5
Research question	5
Objectives of the research	5
Research approach	6
Motivation	6
Why a triple domain?	7
Thesis outline	7
<b>CHAPTER 2</b>	<b>8</b>
<i>Literature review</i>	<b>8</b>
Introduction	8
Research questions unpacked	8
A. Interoperability:	8
B. Scalability	12
C. Access control	14
Discussion of key concepts	20
Grid computing	20
How is grid computing classified?	21
Computational grid	21
Classification based on topology	22
Cluster	22
Security of the grid	23
Security requirements in a grid environment	24
Sensitive information	24
Authorization	24
Authentication and access control	24
Data confidentiality	25
Data integrity	27
Non-repudiation	27

Privacy	28
Availability of resources	28
Authentication	28
Attacks on a grid system	28
Architecture	32
Prototype	33
Computational grid	33
Security model	33
Triple-domain grid-based environment	34
Conclusion	34
CHAPTER 3	35
<i>Research design and methodology</i>	35
Introduction	35
Research design	35
Epistemology	35
Theoretical perspective	36
Methodology	37
Methods	39
Content analysis	39
Architectural framework and algorithm design	39
Simulation and experimentation	40
Application of ssm to the research problem	41
Case study: federated database	50
Conclusion	53
CHAPTER 4	55
<b>RESULTS</b>	55
Introduction	55
SSM CYCLE 1: Content analysis	55
A: Real-world situation of concern	55
B: Purposeful activity: content analysis	55
D: Action to improve the situation	58
SSM CYCLE 2: Architectural design	60
Figure 20: Designing the architectural framework and algorithm	60
A: Real-world situation of concern	60
B: Purposeful activity: architectural design	61
Operational overview of the model	69
C: Model comparison and appraisal	70
D: Action to improve the model	70
SSM CYCLE 3: Simulation	72
A: Situation of concern	72
B: Model comparison and appraisal	72
C: Action to improve the situation	73
Third stage of implementation: EFFICIENT ACCESS CONTROL	91
SSM cycle 4: Case study	91
A: Situation of concern	91
B: Alternative way of solution	92
Terms and definitions as used in this context	94
Role and services specification for DOMAIN A	96
Role and services specification for DOMAIN B	97
Role and services specification for DOMAIN C	98
C: Comparison and appraisal	102
D: How to improve the situation	102
Conclusion	102
CHAPTER 5	104

<i>FINDINGS: interpretation of the results</i>	104
Introduction	104
Research questions revisited	104
a. Interoperability	105
b. Scalability	105
c. Access control	106
Review of research goals.	106
Findings of the content analysis with SSM approach.	106
Findings of the designed framework and algorithms.	107
Findings of the simulation	107
Motivation for adopting SSM research methodology	107
Conclusion	108
CHAPTER 6	109
<i>Discussion and conclusion</i>	109
The need for research contribution	109
Scope and limitations of the research	113
Recommendations and future work	114
<b>Bibliography</b>	<b>116</b>
<b>Appendices</b>	<b>128</b>
<i>Appendix A:</i>	128
Findings of the content analysis	128
Access Control List (ACL)	128
Implementation	128
Strengths of ACL	129
Weaknesses of ACL	130
Mandatory access control (MAC)	130
Strengths of MAC	131
Weaknesses of MAC	132
Distributed Authorization Model (DAM)	132
Strengths of DAM	133
Weaknesses of DAM	133
Context Based Security Model (CBSM)	136
Strengths of CBSM	138
Weaknesses of CBSM	139
Role based access control (RBAC)	139
Strengths of the RBAC model	140
Weaknesses of the RBAC model	141
RBAC Framework model	143
<i>Appendix B</i>	144
Tables containing information for databases of university, hospital and banking	144
<i>Appendix C</i>	147
Report of cross domain query of Figure 41 for domains A and B	147
Cross-domain queries reports	149
<i>Appendix D</i>	151
Report on access control	151
<i>Appendix E</i>	155
Sample schemas for describing interoperability using a template presented in Figure 16.	155
<i>Appendix F</i>	158
<i>Appendix G</i>	164
Sample source code for implementing access control	164
<b>Publications and conference proceeding</b>	<b>169</b>

## LIST OF FIGURES

Number		Page
Figure 1:	Multi-domain policy interaction in a grid-based environment (Haidar, 2002) _____	9
Figure 2:	Grid classification based on topology _____	23
Figure 3:	The process of symmetric encryption _____	25
Figure 4:	Asymmetric encryption _____	27
Figure 5:	Release of message contents (passive attack) _____	29
Figure 6:	Traffic analysis _____	30
Figure 7:	Four elements of the research process (Crotty, 1998, p. 37) _____	36
Figure 8:	The basic shape of SSM (Checkland, 2000, p. 23) _____	38
Figure 9:	The four basic research approaches used in this thesis _____	38
Figure 10:	Simulation cycle using G3S _____	40
Figure 11:	Application of the SSM, adapted from (Checkland, 2000, p. 23) _____	41
Figure 12:	Graphics User Interface (GUI) of G3S _____	44
Figure 13:	Configuring security services and policies _____	45
Figure 14:	Schematic representation of 3DGBE interoperability operation flow _____	50
Figure 15:	Gridification of the 3DGBE database architecture for interoperability _____	51
Figure 16:	Federated database approach of achieving interoperability in a 3DGBE _____	52
Figure 17:	Content analysis method of identified grid challenges with SSM _____	56
Figure 18:	Basic principle of an access control model _____	57
Figure 19:	Adoption rate of the five security models. 2005–2012 _____	59
Figure 20:	Designing the architectural framework and algorithm _____	60
Figure 21:	Phases involved in the proposed architecture _____	61
Figure 22:	STAGE 1 of 3DGBE architectural framework of the proposed model _____	63
Figure 23:	Information flow sequence of the architecture of Figure 22 _____	64
Figure 24:	A 3DGBE stage 2 architectural framework information flow process of Figure 25 _____	69
Figure 25:	STAGE 2 of 3DGBE architectural framework of the proposed model _____	71
Figure 26:	Simulation of designed framework and algorithms _____	72
Figure 27:	Number of available resources in two access control policies (3DGDE and MAC) _____	74
Figure 28:	Secure rate comparison using two approaches _____	75
Figure 29:	Average turnaround time versus number of grid nodes _____	77
Figure 30:	Average turnaround time versus number of service requesters _____	78
Figure 31:	Throughput (MB/s) vs number of nodes _____	79
Figure 32:	Decreased effect of grid nodes on throughput _____	80
Figure 33:	Tri-middleware integration based infrastructure for 3DGBE interoperability _____	82
Figure 34:	Comparative evaluation of interoperability of 3DGBE with the existing system _____	87
Figure 35:	Query for aggregating data from Domains A and B _____	89
Figure 36:	Query for aggregating data from Domains A and C _____	89
Figure 37:	Cross-domain query for joining data from Domains A, B and C _____	90
Figure 38:	A case study approach using SSM _____	91
Figure 39:	Implementation of hierarchical RBAC in 3DGBE _____	93
Figure 40:	Description of hierarchical RBAC access control policy for 3DGBE _____	96
Figure 41:	Grid user, roles and services relationship of hierarchical RBAC for 3DGBE _____	99
Figure 42:	Multiple-domain resource sharing environment _____	129

Figure 43:	Process of identity mapping using Access Control List _____	135
Figure 44:	Process of identity mapping using role based access control _____	136
Figure 45:	Context-based Security Policy (Pretschner, Hilty, Schutz, Schaefer, & Walter, 2008). ____	137
Figure 46:	User-Role-Permission mapping in a flexible RBAC _____	141
Figure 47:	Relationship of elements in a flexible RBAC _____	142
Figure 48:	Grid user administrative task options _____	151
Figure 49:	Available roles and services _____	152
Figure 50:	Grid users' roles, domain hierarchy and their access status. _____	153
Figure 51:	Information about the users, domain and role _____	154
Figure 52:	Sample schema for Domain A (University) _____	155
Figure 53:	Sample schema for Domain B (Hospital) _____	156
Figure 54:	Sample schema for Domain C (Banking) _____	156
Figure 55:	Showing values of 3.68 at $F_{0.05, 2, 15}$ _____	158
Figure 56:	Showing F-Distribution table for 3.68 _____	160
Figure 57:	Showing values of 3.89 at $F_{0.05, 2, 12}$ _____	163





## LIST OF TABLES

Number		Page
Table 1:	Grid simulators with their respective applications _____	42
Table 2:	Summarized features of five security models _____	58
Table 3:	Functions of the architectural components _____	65
Table 4:	Simulation parameters with their corresponding values _____	75
Table 5:	Simulation parameters for $\lambda_1$ , DSR, Rep for Domains A, B, and C _____	76
Table 6:	Comparative evaluation of interoperability of 3DGBE with other models _____	86
Table 7:	Multiple inheritance of permission in a 3DGBE _____	93
Table 8:	Summary of research contributions _____	112
Table 9:	Description of elements in a flexible rbac _____	142
Table 10:	RBAC Framework model _____	143
Table 11:	University database for domain A _____	144
Table 12:	University database for domain A (Continued from Table 10) _____	144
Table 13:	University database for domain A (Continued from table 11) _____	145
Table 14:	University database for domain A (Continued from Table 12) _____	145
Table 15:	Hospital database for domain B _____	145
Table 16:	Hospital database for domain B (Continued from Table 14) _____	146
Table 17:	Hospital database for domain B (Continued from Table 15) _____	146
Table 18:	Banking database for domain C _____	146
Table 19:	Banking database for domain C (Continued from Table 17) _____	146
Table 20:	The values for average Turnaround time and the No of grid nodes for Domains A, B, C _____	159
Table 21:	The values for average Turnaround time and the No of grid service requesters for Domains A, B, C _____	161

## LIST OF PUBLICATIONS

**N.A. Azeez**, and I.M. Venter. "Towards ensuring scalability, interoperability and efficient access control in a multi-domain grid-based environment" Africa Research Journal (ARJ) of the South African Institute for Electrical Engineers (SAIEE), Vol. 104 No 2, 54–68 June 2013.

**N.A. Azeez**, I.M. Venter and A.S. Oyewole. (2012). "Optimizing grid and cloud computing infrastructures to handle security and access control challenges" Annual Conference on World Wide Web Applications, ZAWWW 2012, Mangosuthu University of Technology (Umlazi, Durban), Durban, South Africa, 7 – 9 November 2012

**N.A. Azeez**, and I.M Venter (2012) "Towards achieving, scalability and interoperability in a Triple-Domain Grid-Based Environment" 11th Annual Information Security South Africa Conference ISSA 2012 , Johannesburg , South Africa, 15 – 17 August 2012 (Paper 26) (ISBN-978-1-4673-2158-7, IEEE Catalog Number: CFP1266I-CDR)

**N.A. Azeez**, I.M. Venter and T. Iyamu (2011)"Grid Security Loopholes with proposed countermeasures" , 26th International Symposium on Computer and Information Sciences 26-28 September 2011, Imperial College, London, UK. Grid Security Loopholes with proposed countermeasures, Springer Verlag, London.

**N.A Azeez**, T. Iyamu, I.M. Venter, O.F.W Onifade and R.A. Azeez (2011) "Grid Security: Evaluation of Active and Passive Attacks with Proposed Countermeasures" Science Alert: Research Journal of Information Technology 3(3): 181-190, 2011, New York, U.S.A. URL for online version: <http://scialert.net/qredirect.php?doi=rjit.0000.29742.29742&linkid=pdf>

**N.A. Azeez**, I.M. Venter and T. Iyamu (2011) "A Model for Securing Sensitive Information on a Grid" , SAICSIT (South African Institute for Computer Scientists and Information Technologists) 2011 Postgraduate symposium, 3rd-5th October 2011 at Pavilion Conference Center, Cape Town Waterfront. ACM Postgraduate and Conference Proceedings, 2011. url: <http://dl.cs.uct.ac.za/node/147>.

**N.A. Azeez**, I.M. Venter and T. Iyamu (2011) "Peer-to-Peer Computing and Grid Computing: Towards a Better Understanding" The Pacific Journal of Science and Technology, USA. Volume 12, Number 1. May 2011 (Spring): page 270-276. URL for online version: [http://www.akamaiuniversity.us/PJST12\\_1\\_270.pdf](http://www.akamaiuniversity.us/PJST12_1_270.pdf).

**N.A. Azeez**, I.M. Venter and T. Iyamu (2011) "Grid Computing with Alchemi: An Appraisal and Research Challenges" GESJ: Computer Sciences and Telecommunications, 2011, No.3 (32), pp. 39-43. url for online version:<http://gesj.internet-academy.org.ge/download.php?id=1853.pdf>.

**Article accepted for publication**

**N.A. Azeez**, I.M. Venter and T. Iyamu and A.S. Oyewole “An Investigation into Grid Security Models: Implementation, Strengths and Weaknesses, case study of five Security Models”. Accepted for publication in International Journal of Applied Computing.

**Oral presentations**

**N.A. Azeez**, I.M. Venter and T. Iyamu (2011) “Towards a reliable Model for secured resource sharing in a Grid based environment” Design, Development, Research and Postgraduate Research Colloquium Conference 2011, 26-27 September at 80 Roeland Street, Cape Town.

**N.A. Azeez**, and I.M. Venter. (2012) “Towards ensuring scalability, interoperability and efficient access control in a multi-domain grid-based environment” Post-graduate research open day, New Science Auditorium, University of the Western Cape, South Africa, South Africa, 31st October, 2012.



## ACKNOWLEDGMENTS

The author wishes to express sincere gratitude to Almighty Allah (S.W.T), the Lord of the universe, for providing me with this opportunity to further my career. It is of no doubt that I am indebted in no small measure to my able supervisors in persons of Prof. Tiko Iyamu and Prof. Isabella M. Venter for their academic guide, moral support as well as financial assistance rendered to me in the course of my studies. They have not only supervised my Ph.D but have also guided me in my life. I also wish to express my appreciation to the Department of Computer Science under the able leadership of Prof. IM Venter for financial assistance.

My Parents (late Pa Azeez Aremu Olukokun and late Mrs. Sideeqat Abebi Azeez-Aminu); imparted in me, the spirit of aiming high towards achieving excellence and greatness through hard work and perseverance right from childhood. It is through their training and motivation that I imbibed the attitude of persistence, courage, commitment and determination to ensure and attain success despite challenges. It is on this note I wish to acknowledge their unparalleled efforts towards enrolling for various academic programmes.

It is important at this juncture to express my heart-feelings to my deceased Brother and Sister. Late Alhaji Abdul Waahid Adekomi Ayinla Azeez departed this world when I was about commencing my programme while Alhaja Simbiat Omolara Bello also died few months before the end of my programme after a protracted illness. They both stood in place of my parents and contributed immensely to my success. May the Almighty Allah blot out their sins and admit them into *Al-Janatul-Firdaus*.

It will amount to ingratitude if I fail to acknowledge my Sisters and Brothers who have been painstakingly supportive both morally and spiritually since the inception of the programme.

I sincerely express my gratitude to Sister Haadiyhat, Alhaja Fausat, Brother Qasim, and Monsurat Jadeshola for their moral supports throughout the programme.

It is worthy of mentioning friends in the Department of Computer Science at the University of the Western Cape , most especially the *Monthly Duba Forum* (which serves as a spiritual pillar for seeking Allah's guide and blessings), Mr. Abidoeye Phillip, Mr. Agbele Kehinde, Mr. Rufai Odutayo Raji and Mr Ademola Adesina for their kindheartedness. I appreciate the prayers and moral supports of the entire members of Balogun's family (my in-laws) during the tough period of actualizing this dream.

A word of thank to all the Department of Computer Science staff, most importantly those who offered me invaluable assistance during my course of research. Specifically, I need to mention Mr Reg Dodds for the numerous occasions he assisted me. I wish at this juncture to acknowledge Professors. Henry Nyongesa and Bill Tucker for some of their valuable and constructive criticisms during various departmental postgraduate presentations. Also to be thanked is Ms Rene Abbott, Mr Daniel Leenderts and Ms Fatima Jacobs who assisted in no small measure during my studies. A special word of thanks goes to Kabiesi, Oba Munirudeen Adesola Lawal, (Laminisa I), for providing me a *real-time* spiritual backing when the going was tough and for giving me moral support after the demise of my brother and sister.

Grid Security Services Simulator (G3S) is a simulator which is not open-sourced. Special appreciation goes to the duo of Syed Naqvi and Michel Riguidel from Ecole Nationale Suprieur des Telecommunications (ENST), Paris, France, for providing this simulator free of charge for the purpose of my research.

The staff of the Center for High Performance Computing (CHPC) in Cape Town, who assisted me during training and experimentation, I would like to especially thank Dr. Daniel Moeketsi and Mrs Dorah Thobye for their invaluable assistance.

A heartfelt gratitude to my cousin, Dr. Sefiu Saheed Adekilekun, (who can be regarded as my study-abroad motivator) for his support and motivation. I thank him for motivation and encouragement he gave me to come and study in the Republic of South Africa.

Last but not the least, leaving behind wife and children in the course of study is challenging when in a foreign country. I really appreciate the understanding and incisiveness exhibited by my wife, Muhibat Bisade Azeez and my children, Abdul Waahid Adekomi Alarape Azeez and Sideeqat Iyabo Oluwaremilekun Mojisola Azeez during the course of my studies. May the Almighty Allah (S.W.T) be pleased with you all.



## DEDICATION

This PhD work is dedicated to my brother and sister, the late Alhaji Abdul Waahid Adekomi Ayinla Azeez and the late Alhaja Simbiat Omolara Bello, for their love and efforts towards seeing me achieving greatness in life. May Allah (S.W.T) forgive their sins and admit them into His paradise.



## DECLARATION

I declare that *Towards ensuring scalability, interoperability and efficient access control in a triple domain grid based environment* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Full Name: **Nureni Ayofe Azeez**

Signature:



Date: 14th November 2012



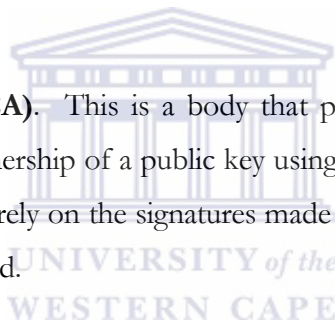


## GLOSSARY

**Architecture.** This is a structural representation and design framework of how a model works. Architecture in this context implies corresponding flow of information with regards to the way model explicitly spells it out.

**Access control list (ACL).** This is a model which contains a comprehensive list of permission attached to an object. It clearly specifies which system or user processes is permitted access to various objects with the corresponding operations that are allowed on the objects.

**Certification authority (CA).** This is a body that provides digital certificates. A digital certificate confirms the ownership of a public key using the named subject of the certificate to allow others believe and rely on the signatures made by the private key which tallies with the public key that is certified.



**Community authorization service (CAS).** This is used when the domains involved, have various capabilities for policy enforcement. In this case, CAS finds a solution to the problem by delegating some authority while taking total control over the provider's resources (Pearlman, Kesselman, Welch, Foster, & Tuecke, 2003). It is a security policy in grids which provides dynamic and scalable mechanisms for authorization in order to monitor and ensure absolute controlled participation of users to distributed resources. It enforces fine-grained access monitoring (Sahai, Graupner, Machiraju, & Moorsel, 2003) procedures in any virtual organization (Lin, Kang Neo, Liang, Guangbin, & Gray, 2007). The CAS server is used by a user accessing community resources which provides rights to the user with respect to the request as well as the role of the user in the community (Pearlman, Welch, Foster, Kesselman, & Tuecke, 2010).

**Community security policy (CSP).** A security measure formulated to guide against the misuse of a computer. It curbs offences ranging from revealing confidential information, accessing inappropriate sites and sending and receiving suspicious and criminal mails.

**Computational grid.** A computational Grid is a form of grids system which allows the connection or the network of various hardware and software resources from various organizations with a mean to providing sharable resources across a multiple administrative domain and to cater for the users need in a reliable and secure manner. In this type of grid system, resources are usually aggregated (Machiraju, Sahai, & Moorsel, 2003) in order to act as a unified processing resource.

**Context-based access control (CBAC).** This is a security mechanism which can be configured to allow specified UDP and TCP traffic through a firewall whenever a connection is established for any network requesting protection.

**Description logic (DL).** It can be used to model concepts, roles and individuals, as well as their relationships. It belongs to the family of formal knowledge representation languages and it is characterized by more expressive power than propositional logic (Franz, Ian, & Ulrike, 2007).

**Distributed computing environment (DCE).** DCE is a software system developed to supply a framework and toolkit for developing client/server applications. DCE provides a dynamic, flexible and scalable distributed environment that resolves complex heterogeneous and network environment problems.

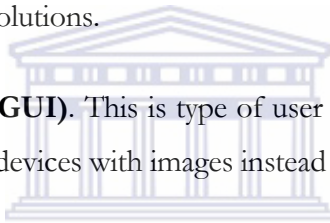
**Discretionary access control (DAC).** This is a type of access control security model designed for restricting access to various objects based on the identity of subjects or groups which they belong. With this model, a subject with certain access privilege can pass such privilege or permission to any subject.

**First order logic (FOL).** It is used in Computer Science, Mathematics and Philosophy by scientists and engineers for modeling. In some cases, FOL is given various names such as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic (Franz, Ian, & Ulrike, 2007).

**Grid security infrastructure (GSI).** This is otherwise known as the Globus Security Infrastructure. It is used for specifying secret and a proxy-based communication between software in a grid based computing environment. The GSI adopts public key cryptography as the background for its functionality.

**Grid security services simulator (G3S).** This is a simulation tool developed for designing and analyzing grid security solutions.

**Graphical user interface (GUI).** This is type of user interface that permits users to relate and interact with electronic devices with images instead of text commands.



**Host identity protocol (HIP).** This is a technology which provides a mean of separating the locator roles of internet protocol (IP) addresses and end-point identifier. It provides a new Host Identity (HI) that depends solely on public keys where end-point identifiers are derived.

**International telecommunication union (ITU).** This is an agency in the United Nations saddles with various responsibilities in information and communication technologies such as maintaining and developing standards and definitions, coordinating and improving statistical ICT measurements on a global level and improving telecommunication infrastructures in developing countries of the world.

**Mandatory access control (MAC).** This is enforced by the operating system or security kernel. Security labels or classifications are assigned to system resources which enable access to processes, devices and people with different levels of authorization.

**Middleware.** These are open source softwares that provide access for a grid user to effortlessly share, distribute, and aggregate the computing resources on a grid system across multiple administrative domains. Without middlewares, sharing and distribution of resources cannot take place on the grid system. Some of the examples of grid middlewares are PUNCH, GLOBUS, OGSA-DAI, Alchemi, Legion and gLite (Gridbus, 2009).

**Multiple levels of security (MLS).** This is a model which permits a computer system to process information at various levels of sensitivities. It allows simultaneous access by different users with different level of security clearance and prevent users from accessing information which they do lack authorisation.

**Northwest Indiana computational grid (NWICG).** This is a partnership which comprises of different scholars, researchers at various universities in the United States. They established a cyber-infrastructure that provides solutions to various complex scientific problems (Purdue, 2011).

**Peer-to-peer (P2P).** Peer-to-peer can be referred to as a communication structure where each computer (node) involved has equal abilities and capabilities and any computer (node) can establish a communication process. In most cases, peer-to-peer communications is established by providing client-server capabilities with all the nodes connected together (Verma, 2004).

**Personal digital assistant (PDA).** It is any handheld device or equipment which comprises of telephone, computing, fax, networking features and the internet. PDA can carry out the function as a fax sender and receiver, mobile phone personal organiser and web browser.

**Public key infrastructure (PKI).** It can be defined as a combination of software, hardware, policies as well as procedures required to control, create, administer, supervise, store and withdraw digital certificates. In security, a PKI binds public keys with each user identities with the aid of a certificate authority (Price, 2003).

**Rivest, Shamir and Adelman (RSA).** A popular public-key encryption algorithm.

**Role-based access control (RBAC).** In computer science security, RBAC is a security model for allowing access to genuine and authorized users alone. With RBAC, security procedures are greatly and perfectly organized and simplified by making use of roles, hierarchies, and constraints to organize privileges ( Zhao & Chadwick, 2008).

**Security model (SM).** This is a statement that states the requirements are crucial to adequately support and implement a certain security policy.

**Sensitive information (SI).** This is information that should be accessed, viewed and updated only by the authorized user. It is considered to be sensitive because it is not meant for the general public but rather for a specified class of users who have authority to do so. If an unauthorized person can access the information it renders the information's sensitivity and confidentiality worthless and useless.

**Simple public key infrastructure (SPKI).** It was developed only to overcome the problems that emanates from the over complication and scalability of traditional X.509 public key infrastructure. SPKI is a framework to support security for a wide range of internet applications.

**Soft system methodology (SSM).** It was conceived by Peter Checkland of Lancaster University as a problem solving approach to system development (Checkland, 2000). This methodology is suitable for research problems with no obvious or clearly defined solution reference and is thus appropriate for this research.

**Trusted third party (TTP).** In cryptography, a trusted third party (TTP) could be regarded to as an entity that easily facilitates mutual interactions between two parties. The two parties involved make use of trust reposed in TTP to secure their interactions and communications.

**Web content management system (WCMS).** It is a software system which allows creation and management of websites with relative ease.

## *Chapter 1*

### STATEMENT AND ANALYSIS OF THE PROBLEM

#### **SKETCHING THE BACKGROUND**

In this chapter, the background information of the study is provided to contextualize the research process. The chapter is divided into seven main sections. The first section provides the definition of a grid in the context of this study, while the second section explores the rationale for securing a grid. In the third section, an explanation of the challenges being addressed is provided and the motivation for addressing those challenges is provided in section four. In the fifth section, the research problem is stated and unpacked into two research questions. In the seventh section, the outline of the thesis is given.

#### **Definition of a grid**

The concept of a grid system is analogous to a water grid system. The facilities of a water grid system make it possible for individuals and groups to open a tap and collect water without knowing or understanding exactly where and how such water is being processed (Buyya, 2002). Grid computing provides endless and ubiquitous access to expensive but high-quality computing resources to users wherever they find themselves and without them having to know exactly where the data is being processed.

It was in the late 1980s and early 1990s that researchers across the world began to explore various types of distributed resources on the Internet. Some of these researchers gathered and optimized thousands of workstations for various parallel applications such as graphics rendering, simulation and molecular design. Others teamed up to connect various supercomputers into a virtual metacomputer over a wide-area network (Gentzsch, 2000) in order to understand the actual benefits of a distributed environment and to enhance and develop it further.

The term “grid computing” was coined by the duo of Ian Foster and Carl Kesselman (Foster & Kesselman, 1997), who organized a workshop entitled “Building a Computational Grid” (NHSE, 2009) at Argonne National Laboratory, The purpose of the workshop was to address challenges which were related to distributed computing and its applications.

Since the introduction of the term “grid computing” there have been many definitions of it. One of the definitions includes Buyya (2002), who defined a grid as follows:

*The "grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of resources distributed across multiple administrative domains based on their (resources) availability, capability, performance, cost, and users' quality-of-service requirements." (Buyya, 2002, p. 35).*

The South African Grid (SAGrid) is a typical example of a functional grid where a conglomerate of tertiary institutions (universities, laboratories and the Meraka Institute) are collaborating in the sharing of resources (Grid, 2010). The five universities that are currently on the SAGrid are the University of Cape Town, University of Pretoria, University of the Free State, University of Johannesburg and University of the Witwatersrand (GStat, 2010).

### **Why secure the grid?**

The need for preventing sensitive and important information from being copied, altered, divulged to unauthorized users or manipulated has brought about the need for security on a grid system ( Hwang & Yang, 2010).

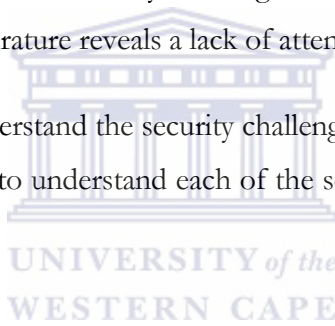
Over the years, there have been an increasing number of applications, data and information on grids. According to Mohteshim (2005), this has in turn increased the need for greater dependability and efficiency of the grid environment. Dependability and efficiency imply security. Without security a grid cannot be considered to be dependable. However, security models on the grid are difficult to implement and sustain, due to the complexity of the grid environment. Traditional access-based control models are based on recognized inadequacies ( Sandhu, Coyne, Feinstein, & Youman, 1996) and there is thus a need to replace them with more flexible models which are relevant to distributed environments.

### **Security challenges to grid computing**

As in a many environments, grid computing continues to encounter security challenges. In order to address the security challenges within the grid-based environment, the need for understanding is critical. Some of the security challenges include authentication, manageability, coordination, interoperability, scalability, and efficient access control.

The present study focused on three main challenges, i.e. interoperability, scalability, and efficient access control. This is primarily because they are considered fundamental, prevalent and significant to the effective utilization of grid computing (Bouwman, Mauw, Eindhoven, & Petkovic, 2008). Most importantly, previous studies have not addressed the three issues (components of the security challenges facing grid computing) simultaneously and more so, a survey of literature reveals a lack of attention to these three basic challenges.

It was first necessary to understand the security challenges in order to achieve the objectives of the study. In an attempt to understand each of the security challenges they are explained in more detail below.



#### *Scalability*

Scalability is the ability of a grid system to act efficiently, sufficiently and adequately in handling both a small and large number of nodes. This includes the processing capacity configuration in order to achieve a secured multi-organisation resource-sharing environment (Detsch, Gaspar, Barcellos, & Cavalheiro, 2004).

The major challenge of grid computing lies in the flexibility to scale up the number of users and resources from different domains (Rahmeh & Johnson, 2008). The essence of scalability is to purposely cater for future expansion (Ramaswamy, Liu, & Iyengar, 2007). It becomes a security challenge when access control is a bottleneck in the grid infrastructure and cannot scale, especially when the number of grid domains and users increases. For a grid environment to be scalable, therefore, there is need for a centralized administration system as well as regular updates of related security policies (Al-Bayatt, Zedan, & Siewe, 2009).



*Interoperability*

The interoperability component provides the ability of various systems to operate within the grid. This is in order for them to exchange, share and utilize information across different platforms (Abbes, Barbera, Jemn, & Mazigh, 2008). Interoperability is considered a major security challenge due to factors such as disparate and unequal security policies, as well as the incompatibility of access requirements from different environments (Shen, Yang, Tian, & Liu, 2006).

Diversity in the various security policies poses a challenge to attempts to access various domains (Shafiq, Joshi, Bertino, & Ghafoor, 2005). According to Goyal(2011), the characteristics of an interoperable grid-based environment include:

- the presence of a central authority for security and trust;
- heterogeneous resources, service discovery and management; and
- the interdependence of security infrastructures.

*Efficient access control:*

Efficient access control (EAC) is intended to enforce control over who can interact with resources on the grid network. EAC can be achieved through different means such as authentication and authorisation with the aid of an appropriate access control model.

EAC remains a challenge in grid computing mainly because a very large number of users are involved. The users are often considered to be dynamic in their requests (Snelling, Berghe, & Li, 2004). This could be attributed to the fact that each domain on the grid has its own policies and the various domains are autonomous (Lang, Foster, Siebenlist, Ananthakrishnan, & Freeman, 1990).

To achieve EAC, regulation is required. To regulate and find solutions to the factors that impact the EAC within the grid platform, a role-based access control model is proposed. This model has role assignment, role authorization and domain hierarchy as three primary rules.

Based on the focus of the study as presented above, the motivation for the study is discussed in the next section.

## **RESEARCH FRAMEWORK**

### **Research question**

To investigate the challenges posed by scalability, interoperability and efficient access control of a grid, the following research question was articulated:

How should information on a grid be secured without compromising the accessibility and availability of resources?

The research question was further divided into two sub-questions as follows:

- How should interoperability and scalability be put in place to ensure the optimal utilization of grid computing?
- How should efficient access control be administered for the entities on the grid in order to monitor and control permission to access various resources?

### **Objectives of the research**

The objectives of the research study were two-fold:

- (i) To investigate various models relating to scalability, interoperability and access control on grid computing systems, i.e. how they are implemented, their mode of operations, and their strengths and weaknesses.
- (ii) Based on the findings from the investigation, to develop architectural models. The models are aimed at providing scalable, interoperable and efficient access control on the grid.

**Research approach**

In order to achieve the objectives of the research, the following approaches were adopted:

Objectivism was the epistemological stance of the research and positivism was the theoretical perspective. The methodology that was used to manage the research was Soft Systems Methodology (SSM) and the methods that were adopted were:

- a) Content analysis: analysis of related literature;
- b) Architecture design, based on the dynamic role-based access control (RBAC) model;
- c) Using simulation and coding to measure the performance of the architecture model; and
- d) Quantitative analysis of the simulation findings.

**Motivation**

Scalability, interoperability and efficient access control have been identified as the greatest challenges facing grid computing, since sensitive, classified, expensive, confidential and valuable resources are shared between various organizations ( Lu, Cao, Chai, & Liang, 2008). Other challenges for grid computing are management, scheduling, coordination and resource sharing (Foster, Yong, Raicu, & Lu, 2008). The former challenges remain greatest obstacles towards actualizing the dream of exploiting the full benefit of a grid. A grid-based environment that is not scalable constrains and limits the number of entities (resources and users) on the grid while a non-interoperable grid-based environment reduces the level of application from platform to platform. Finally, a grid with a poor access control framework poses a great security challenge, since information should be accessed only by authorized subjects hence the need for appropriate access control mechanism to achieve this ( Sandhu, Coyne, Feinstein, & Youman, 1996).

To overcome the problem of scalability, interoperability and poor access control, a triple-domain grid-based environment (3DGBE) architectural framework was developed based on a flexible role-based access control (RBAC).

Some of the traditional access control mechanisms currently being adopted in grid computing such as discretionary access control (DAC), mandatory access control (MAC), access control list (ACL), capability-based access control (CBAC) and multi-level security (MLS) are identity based, which makes them neither interoperable nor scalable and thus they cannot accommodate the increasing number of users or allow a range of policies on their platform ( Hu, Ferraiolo, & Kuhn, 2006).

### **Why a triple domain?**

The choice of a 3DGBE is basically to create and establish competition among the entities (users and resources) available on the grid. A 3DGBE which is designed with the local security monitoring unit (LSMU) from each of the domains, interacts directly with the central security monitoring unit (CSMU) of the architecture with the intention of sharing and requesting resources. The LSMU gives authorization at local level to any user willing to access resources. The CSMU, on the other hand, receives messages from the LSMU indicating authorisation to a grid user locally.



### **THESIS OUTLINE**

In Chapter 2, the literature related to this study is reviewed. Various categories of grid system and grid classification (based on its topology) are discussed, as well as the various security policies in a grid-based environment. Furthermore, various categories of security attacks are analysed and the chapter concludes with a detailed overview of grid middleware.

A comprehensive overview of the research design and methodology is presented in Chapter 3. An overview of the prototype used in the research effort as well as the algorithm that underpins it is discussed. The Grid Security Services Simulator (G3S) that was used for evaluating the performance of the architecture and algorithm is discussed.

In Chapter 4, the architectural models and algorithm are presented along with results obtained, followed by a discussion around the architectures presented in Chapter 5. Finally, in Chapter 6, conclusions are drawn and recommendations made for future work.

## *Chapter 2*

### LITERATURE REVIEW

#### **INTRODUCTION**

In the previous chapter, the problem was stated and the research background was sketched. The chapter was concluded by stating the objectives of the research and presenting the research approach.

This chapter presents a review of literature related to the study. The focus is on the application of grid computing and the research work that addresses issues related to access control, interoperability and scalability on the grid. The chapter is divided into two sections. The first covers the review of literature based on the research questions posed in Chapter 1, and thereafter addresses scalability, interoperability and access control, while the second presents a review of the key concepts in this field of study.

#### **RESEARCH QUESTIONS UNPACKED**

**How should interoperability and scalability be put in place to ensure the optimal utilization of grid computing?**

##### *A. Interoperability:*

The research of Tari and Fry(2001), considered interoperability as an important metric in terms of the accessibility of information. They adopted policy aggregation to realize an interoperable grid-based environment.

By definition, security policy aggregation is a process of bringing various security policies together from various domains in order to achieve a conflict-free resource-sharing virtual environment (see Figure 1).

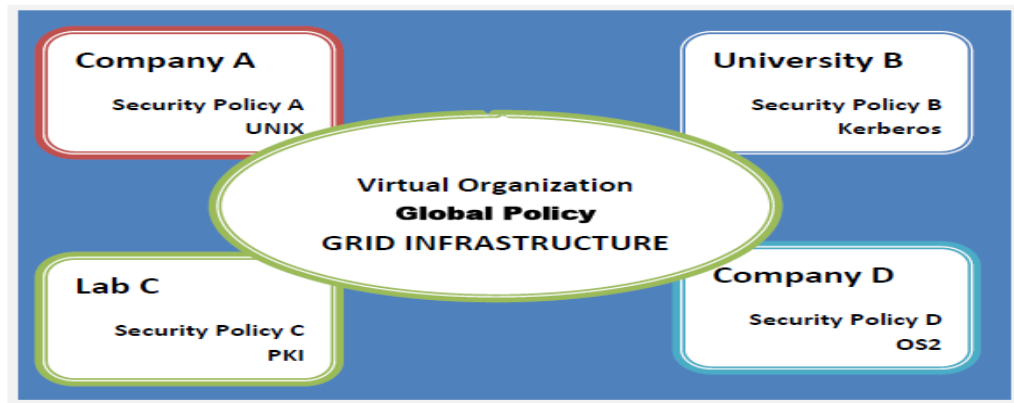


Figure 1: Multi-domain policy interaction in a grid-based environment (Haidar, 2002)

In Figure 1, Company A uses policy A on a UNIX platform, University B uses policy B on a Kerberos platform, Laboratory C uses policy C on the PKI platform, and Company D uses policy D on the OS2 platform. The “global policy” is an overarching policy for all domains on the grid. Security policy aggregation is when various local security policies are reconciled with global policies in order to access grid infrastructure. This is essential in order to resolve conflict between global (Lupu & Sloman, 1999) and local policies (Haidar, 2002).

To ensure interoperability across different domains, Tari and Fry, (2001) proposed global access control (GAC) to enforce an aggregated security policy. They used a distributed object kernel security service to enforce and aggregate local and general security policies on the grid. In order to allow the control of data aggregation, they provided the security framework; federated logic language (FELL) and a logic-based language. The security constraint was enforced in their model by mapping state transition graphs (which models different nodes) on the grid. This approach is effective and enforces various security measures, but it is not scalable, since it does not allow more nodes to be added to the existing grid.

Kumari et al. (2010) proposed a collaborative trust enhanced trust security model for an interoperable grid environment. They made use of client node and node registry to ensure security when a client demands certain services. However, the implementation of their work revealed its weaknesses in terms of access control and scalability, as it was difficult to alter

the system in order to accommodate an increasing number of users, computing entities and resources. Also, the proposed model did not provide secure communication among the grid users (Kumari, Shakti Mishra, & Kushwaha, 2010).

In combating the challenge of interoperability in a grid based environments, Wang et al. (2006), developed a modular information provider (MIP) to technically combine diverse sources of grid information services. An MIP requires minimal human intervention for aggregating a large pool of information services ( Wang, Shook, Padmanabhan, Briggs, & Pearlman, 2006) and adopts a modular approach to ensure interoperability. An MIP uses the XML version of GLUE Schema 1.2 and Globus MDS4 for implementation. The modular pattern of MIP is designed to simplify the management and maintenance of grid information by automatically and repeatedly aggregating information sources. However, this framework has limitations in terms of its application on various operating systems. The middleware that can be used is limited only to Globus, which does not extend MIP to capture a MIP-level aggregator to supply information for the whole site.

The need for interoperability for a complex collaborative multi-disciplinary design product has led to the introduction of a semantic grid, which is a combination of grid technology and a semantic web in the work of ZeFeng et al. (2010). The introduction of a hierarchical model in a semantic grid leads to digital system interoperability that is based on a semantic grid ( ZeFeng, HaiCheng, & Rong, 2010). This work is novel in its entirety, but the implementation provided does not prove beyond doubt how a semantic grid can achieve interoperability among digitizing systems.

Having noticed the non-interoperable nature of many existing monitoring systems due to lack of consistency, a generic monitoring framework was proposed by Shen et al. (2006) to address interoperability in a grid-based environment. They designed a consistent data provider that was designed on a GMA producer to achieve uniform information representation across a grid platform. They developed republisher and monitoring service agent to ensure information sharing and aggregation ( Shen, Yan, Qiang, Wu, & Zhang, 2006). Implementation was carried out using Ganglia, PBS and GridView. The approach

adopted in this regard is not suitable for a scalable and flexible grid environment, however, because a pacified number of entities are required.

The challenge arises from lack of interoperability of various grid across the globe was addressed in the work of Abbes et al. (2008). They proposed a super worker node-based architecture which aggregates several services of a gLite-WN and an lcg-CE ( Abbes, Barbera, Jemn, & Mazigh, 2008) with service that manipulates the gLite job to a XtremWeb-CH job. The relevance of this work has not been empirically verified to ascertain its effectiveness.

The use of public key infrastructures (PKIs) to provide maximum flexibility and interoperability in order to permit grid resource use manages to securely provide access to various systems across the multiple administrative domains. In this approach, various services are added as PKIs grow to enhance interoperability and security. Applications and users must strive to confirm the availability of resources especially when the certification authority (CA) is unfamiliar to the reliant party (Pala, Cholia, Rea, & Smith, 2008). The work is able to achieve interoperability through the integration of both grid security infrastructure with PKI resource query protocol (PRQP).

Di et al. (2006) proposed a novel approach for connecting various heterogeneous grid platforms to achieve interoperability. They solved the problem of interoperability using the concept of ontology. They focused on solving the problem of the semantic inconsistency of different grid platforms and by developing a unique ontology. The ontology is transacted with adapted and depicted by a resource description framework (RDF). The algorithm developed was programmed into an adapter that gives faces for platforms' administrators. The approach was tested on two grid platforms ( Di, Jin, Li, Chen, Qi, & Wang, 2007). The approach, although effective, does not give room for flexibility in terms of resources and a platform for optimising the design and algorithms.

Ensuring interoperability through the application of information-flow theory and adopting it to enhance information interoperability in the distributed system, specifically on the grid,



was researched by Weihua and Shixian (2005). The work presents an agent-based information interoperability framework with the layered model. It achieved automatic information interoperability through the aggregation of ontology and agent technology. The weakness of this work lies in its inability to take into account a model for advanced network application.

VCF components established on ESA's current Excel-based concurrent design facility was developed by Watkins, Mark, Leonard, and Surridge (2007) in order to demonstrate the integration and interoperability of grid computing in distributed design. They achieved this by creating a STEP-compliant database that stores and validates design parameters. The database was designed in such a way as to reveal a GRIA 5.1 grid service, while a .NET/WSE3.0 was established to ensure parameter parsing between the STEP database and the Excel client. The policy updates were able to achieve high level of interoperability between various entities on the platform. However, to enhance the performance of the framework, additional WS-Trust token services would be required at various domain sites ( Luo, Ni, & Yong, 2009).

### B. Scalability

Having noted the challenge of scalability in a collaborative virtual environment (CVE) due to limited network technologies and computing resources, Zhang et al. (2007) developed a hybrid (combined) communication model based on open grid services architecture (OGSA). Consequently, a scalable CVE architectural framework established on OGSA was developed. This was achieved by using a virtual world management service (VWMS) to ensure and implement scalable CVE ( Zhang, Wu, & Wu, 2007). They implemented the framework on a GT5.0 to validate the effectiveness of the architecture. However, the architecture has limited application as it is applicable to virtual campus systems only (Grimshaw, 2008).

A scalable authorisation was developed by Wang, Zheng, and Song (2005) in order to achieve scalable grid-based environment. An authorisation framework that was established on a distributed authentication servers, as well as an extension to the information service of

a grid-based system. Scalability is achieved in this framework by developing local authentication servers and incorporating it into the information service, which ensures that various services are enabled at the grid system level and local host level. The flexible nature of the grid system was addressed by adopting an RBAC security model that binds grid users to various assigned roles rather than binding to the resources.

In an attempt to ensure scalability for scientific applications on grid computing with a service-oriented architecture, Sanjeevan et. al, (2005) developed a framework that provides a uniform abstraction for a category of applications and ensures a generic application service to envelop and wrap these applications as services. The novel contribution of this approach lies in the automatic nature of the entire process, which is carried out without the need to request grid system downtime or coding. This approach represents a very useful technique for building grid portals through which a great number of applications can be dynamically established. However, this framework could not implement complex simulation process that involves multiple applications that can be managed with constant user interaction by giving a dependable workflow specification and aligning it with a workflow engine.

Gor et al. (2005) achieved scalability by unified grid management and data architecture (UGanDA). The framework consists of two main components, GridWorm and MAGI, which are the grid workflow manager and grid infrastructure manager, respectively (Gor, et al., 2005).

While GridWorm allows the workflow to be specified and managed with an intuitive interface, MAGI reduces the task of job submission and diligently executes jobs over a grid infrastructure. The synergy between the components allows the users to specify high level requirements along the workflow system, which subsequently uses MAGI's management services to manage the infrastructure (Gor, et al., 2005). The proposed framework sounds interesting, but its application does not expand to developing the data management capabilities across a large grid network.

The challenge of scalability was also addressed in the research carried out by Li (2005). He developed an authorisation framework based on grid authorisation servers. The objective of

his framework was to address and overcome various challenges discovered in existing authorisation architectural designs by providing a scalable authorisation and authentication approach that could meet the requirement of a flexible grid based environment. With this approach, scalability was achieved with the use of distributed servers. The flexible nature of a grid system was handled using RBAC, while authorisation was conveniently adjusted and manipulated as changes arise in the grid system environment (Li , 2005). The results obtained from the experiment were not convincing enough to confirm the efficiency of the proposed framework.

**How should efficient access control be administered for the entities on the grid in order to monitor and control permission to various resources?**

*C. Access control*

Access control in grids involves appraising each request submitted by a virtual organization's entity to access any desired resources in the grid, in a bid to determine if the request by the grid user should be denied or permitted, based on the policy available for accessing resources in the virtual organisation (VO). In summary, controlling access to resources on the grid requires three important phases: authentication of the initiator, an authorization decision for the submitted request and enforcement of the request (Laccetti & Schmid, 2007 ).

Access control permits an authority to monitor access to resources in any computer-based information system or physical facility to protect important, classified, confidential or sensitive information or resources ( Hu, Ferraiolo, & Kuhn, 2006). Access in this model can either be *allowed* or *denied*, depending on the authentication methods. Access control could also be interpreted to mean a process of exercising absolute control over who should use or access a particular resource (Bouwman, Mauw, Eindhoven, & Petkovic, 2008).

In the work of He et al. (2008), a model was developed for the grid based on a public key and double identity authentication to ensure access control among various entities. The model based on the RSA cryptosystem of Euler's theorem ensured both authenticity and

confidentiality (Dallon et al., 2007). The double identity authentication approach included a “time parameter” on the server side using both the server and a client-produced password that changes over time (He, Li, & Hu, 2008). This model was not scalable and dynamic, as provision was not made for the expansion of users.

Attribute based access control (ABAC) systems, such as Akenti and PERMIS (Lang, Foster, Siebenlist, Ananthakrishnan, & Freeman, 1990) have been in use on several grid applications some time. These authorization systems allow only their own rules and therefore do not allow other rules and policies. Dynamic ABAC would be preferred for a dynamic grid computing environment, which is not the case in this ABAC model, where there is no provision made for interoperability across various domains on the grid.

McLean came up with a framework using mandatory access control (MAC) with an architecture that allow for changes in security. He employed algebra to construct his model, which paved the way for discretionary access control for N-persons (McLean, 2008). This model, however, did not provide for the separation of duties as an ingredient for a secure grid environment.

In order to ensure efficient access control in distributed environment, architecture for securing authorship was developed by Ferreira, Berstis, Armstrong, Kendzierski, Neukoetter and Takagi (2003). An architecture to ensure temporal and authorship protection of information was developed which can be used to resolve conflicting claims on intellectual property. These authors applied PKI techniques to institute access control and reliability on a grid. To ensure a reliable environment, a trusted entity using a certificate authority (CA) provides a digital certificate to some entities, which provides information about the key owner and public key. The model uses an encryption on a private key to perform a digital signature.

The challenge of this model is that it is not scalable, as it has limited application and it also reduces the number of users. Also, it does not permit flexibility among the grid users if conflicting roles are granted among various users.

When researching access control, Laccetti. and Schmid (2007) came up with a framework for reliable security that ensured accessibility to resources on the grid based on grid security infrastructures (GSI) and community security policy (CSP) which captured the policies and rules of the grid. A trust relationship, based on a cryptographic key (Menezes, Vanstone, & Van Oorschot, 1997), was used as the guiding principle for this framework. It was found that authentication (implemented at grid levels) develops a trust relationship that is transitive (which is not the case when authentication is used) (Laccetti & Schmid, 2007 ). This model was characterized by a lack of flexibility, as it had limited application (Kumari, Shakti Mishra, & Kushwaha, 2010) , coupled with the fact that it does not give room for scalability , which is addressed in this thesis.

Architecture based on user identity authorization was developed by Junrang et al. (2004). The model allowed a user to get a grid agent and sign a certification to provide inter-domain authentication with a user agent. After this, the user sent an identity authorization request to Globus Secure Infrastructure (GSI). When a user was authorized, a pair of keys was created by GSI, which forward a private key to the user for transferring the public key with the user grid ID to the storage agent. At this stage, the user sent a request already signed with the received private key to the mass storage agent to subsequently verify it with the public key. This model provided a simple method of access control, but it was not reliable when the number of domains was increased (Junrang, Zhaohui, Jianhua, & Mingwang, 2004). Also, there was no policy aggregation among various domains and hence no flexibility among users.

A novel mechanism to provide access control and reliability to achieve flexibility on a grid was developed by Yan Li et al. (2008). The prototype was developed to protect sensitive files and data across domains. In the model, operation, object, subject and policy constraint defined the entire mechanism. To evaluate the new mechanism, comparison between various results was presented with different policies, since the focus was on access control and reliability in order to achieve flexibility. Implementation revealed that access control was suitable for accessing resources without hindrance. It also revealed that the prototype was

reliable. However, the new mechanism was only suitable for a limited grid environment and did not allow for interoperability (Li, Sun, Chen, Ren, & Luo, 2008).

To further the ability to access resources on a grid system, the SESAME prototype was developed by Zhang and Parashar (2006) using a dynamic context-aware access. The model complemented a subsisting authorization mechanism to dynamically grant permissions to various users based on their context. In this model, each user was assigned a role subset from the role universal set. Also, each resource was given a permission subset from a universal permission set to each of the roles that had the right to access the resource (Zhang & Parashar, 2006).

According to the model developed, if user N wished to log into the system with his/her PDA based on his/her credentials, the authorization service assigned the user a set of roles and also set up an access control mechanism on the same PDA that regulated and maintained the role and state of the machine. Based on the level of the user's security wireless connection, a dynamic RBAC policy was used to select an appropriate role. The main concern in the implementation of this model lay in its reliability and policy aggregation towards achieving flexibility so as to ensure accessibility to various resources (Hey & Trefethen, 2002). It was, however, noted that this model was applicable in an organization where the number of roles and permissions is relatively small. Hence, the model was not scalable (Foster, Kesselman, & Tuecke, 2001).

Also addressing this question was Cai, Liu, Guo, Zhang, and Geng (2009) reliable model built on a password-key exchange (GPAKE) The model could satisfy various requests submitted by different users to the authorized access control system alone. However, it was faced with serious attack because the password could be guessed. Also, it provided neither room expansibility in any form nor interoperability in a grid-based environment.

Chen et al. (2006) developed an access control framework which consisted of three basic fundamental elements: meta-model, meta-data and a meta-protocol. The meta-model was used to map relationship among different entities through level architecture while the meta-

level allowed the various computing entities to implement the reflection function of the system. These entities used the meta-protocol in manipulating and retrieving meta-data in the meta-level. The objective of this framework was to manipulate and adjust the system's activities and status based on the relationship between the system's implementation and the user's action (Chen, Wu, Wang, & Chi, 2006). The weakness of this framework lay in its inability to empirically evaluate the framework to affirm and confirm its effectiveness.

The work of Cirio et al.(2007) showed how semantic web ontologies can be adopted and used to establish an efficient access control system. They made use of the RBAC technique and improved it with contextual attributes. The methodology allowed for the flexible allocation and association of roles among users on the grid ( Cirio, Cruz, & Tamassia, 2007). The classification of resources and users, together with access control consistency policies, was carried out with the assistance of a Description Logic (DL) reasoner. They achieved this by reducing the expressive power of DL formalism using SPARQL queries to cleanse the output of the DL reasoner. The weaknesses of this approach lay in the complexity of its implementation and limitations in the area of application.

Jeong et al. (2006) developed an authorisation and authentication architectural framework for a grid-based system. In their design, they made use of SAML (Security Assertion Markup Language) and XACML (eXtensible Access Control Markup Language) for building single sign-on and authorisation (Jeong, Yu, Shin, Shin, Moon, & Lee, 2006). Although, the architecture permitted the possibility of establishment and implementation of various single sign-on techniques for a secured grid computing environment, the approach was static and did not allow variation in policies for the established scenario.

An enhanced RBAC security model was presented in the work of Aziz et al.(2006) They clearly defined a risk approach for the model that expressed various elements of combinatorial, operational and conflict of interest risk in a particular scenario. Their model consisted of various mechanisms for risk reduction such as stack checking, a firewall and redundancy (Aziz, Foley, Herbert , & Swart, 2006). The problem with this framework was

its inability to monitor access efficiently when the number of entities on the grid was increased (Nithya & Banu, 2010). It also lacked flexibility.

Geethakumari et al. (2009) proposed an access control architectural framework for the implementation and adaptation of RBAC in a grid-based environment. They solved the various access control challenges by proposing role equivalence among various domains through mapping of the role at the local domain level with its equivalent global role (Geethakumari, Atul, & Sastry, 2009). The final authorisation was achieved through the mapping of resource access policies with the global role ranking. The challenge in this approach lay in its inconsistency in terms of ranking pattern, which consequently resulted in static role authorisation.

A password-based authorisation mechanism was developed by Cai et al. (2009). An access control framework with a three-party password key exchange can guarantee user requests to the legal and authorised access control. The protocol they developed used existing GPAKE to authenticate the identity and subsequently provide access privilege to the applicant (Cai, Liu, Guo, Zhang, & Geng, 2009). This access control could only provide lower security level (Humphrey, Thompson, & Jackson, 2005) for the system and the danger of attack by hackers guessing the password was very high.

A framework that made use of risk and trust as the two basic vital parameters for access control decision was proposed in the work of Li et al. (2008). The mechanism was intended to represent confidence in the peer and to guarantee the safety of users' information (Li, Sun, Chen, Ren, & Luo, 2008). The work was novel in its entirety, but evaluation models were not introduced to truly refine the evaluation of both risk and trust. Also, the assessment of the amount of input and output precision remained a difficult challenge.

In an attempt to find a solution to the challenge of access control on a grid platform, Chen et al. (2008) adopted a fuzzy approach of trust by applying the fuzzy theory of reasoning and representation to the multi-domain access control approach (Chen, Luo, & Ni, 2008). The case study claimed to achieve effective and efficient access control in a multi-domain grid-



based environment. However, the two-level fuzzy approach presented was cumbersome and ambiguous. The approach was difficult to digest and did not allow the resources to be expanded.

## **DISCUSSION OF KEY CONCEPTS**

Since all research has associated key concepts around which the study is developed, it is necessary to clarify the meaning of such key concepts. In this thesis, the key concepts that will be discussed in this section are grid computing, security, computational grid, security model, authorization, authentication, prototype and architecture.

### **Grid computing**

Grid computing can be regarded as a form of parallel or distributed computing which allows the sharing of resources across various organizations. One of the key factors that distinguishes grid computing from cluster computing is that a grid system tends to be loosely coupled. Aside from this, a computing grid is usually designed to run just a single application at a time and it makes use of a grid middleware (Buyya, 2009), which serves as intermediary between the resources and the machine. There are basically three prominent categorizations of grids, namely: computational grids, data grids and service grids (Alfawair, Aldabbas, Bartels, & Zedan, 2007).

The concept of a grid system is analogous to a water grid system (Chetty & Buyya, 2002). The facilities of water grid system make it possible for anyone to open the tap and fetch water without knowing exactly where such water is being produced or processed. The technology behind computing grid is very similar to this (Jacob & Fukui, 2005). Grid computing therefore provides users with endless and ubiquitous access to expensive but high quality computing resources to the user wherever they find themselves. The concept of grid computing entered the spotlight in the mid-1990s when it was considered as a basic system and structure for solving problems in science and engineering. Since then it has also been incorporated into the process of finding solutions for both scientific and commercial applications (Foster, Kesselman, Carl & 2003).

## **How is grid computing classified?**

### General classification

In terms of their general classification, three main categories of grids can be identified, based on their application (Alfawair, Aldabbas, Bartels, & Zedan, 2007). Each classification has to do with the solution they offer at a particular point in time. The classifications are as follows: computational grids, data grids and service grids.

### **Computational grid**

A computational grid (Foster, Kesselman, Tsudik, & Tuecke, 1998) simply means a connection or network of various hardware and software resources from various organizations with a means to provide sharable resources across a multiple administrative domain and to cater for users' needs in a reliable and secure manner (Rosado, 2009). In this type of grid, resources are usually aggregated (Buyya, Abramson, & Giddy, 2001) in order to act as a unified processing resource. A few of the numerous problems that are solved by computational grid systems are modeling and simulating complex scientific and engineering problems, diagnosing medical conditions, forecasting the weather and many others. The Northwest Indiana Computational Grid (NWICG) is a typical example of a computational grid (NWICG, 2008).

### *Data grid*

As the name implies, the data grid (Foster & Kesselman, 2003) handles and deals with data. It allows data to be shared and transferred among various users on the grid. In other words, a data grid can be regarded as a basis for sharing, managing and controlling very large amounts of data among authorized users distributed across different networks. The need for a data grid is pivotal for data distribution in many organizations. The data grid is an innovation that has been widely accepted and implemented in academe, industry, research institutes and the banking sector. However, significant security challenges are experienced at the highest level of opportunity being realized on data grid system; these challenges are currently being researched by various researchers in both academia and research institutions. The main objective of a data grid is to model the future generation of computing resources

that will be able to handle very large-scale database integration (Alfawair, Aldabbas, Bartels, & Zedan, 2007).

#### *Service grid*

There are three categories under this classification: multimedia service grids, on-demand service grids and collaborative service grids. A multimedia service grid provides services for various real-time multimedia applications such as virtual reality (Buyya, Abramson, & Giddy, 2001), interactive TV and computer games. A collaborative service grid combines different infrastructures to provide new services, while an on-demand service grid provides real-time collaboration and interaction for resource sharing (Buyya R. , 2003).

#### *Classification based on topology*

Grid computing can be classified in terms of topology into four main groups, i.e, clusters, intra-grids, extra-grids and inter-grid ( Foster , Kesselman, Lee, Lindell, Nahrstedt, & Roy, 1999). This is depicted in Figure 2.

#### *Cluster*

This is the smallest form of grid both in scope and size (Alfawair, Aldabbas, Bartels, & Zedan, 2007). It involves a combination of various servers to generate high computing power in comparison to what is obtainable in an offline (standalone) system. This form of grid computing is developed to solve problems in a unit or department and is usually implemented in a university campus intranet.

#### *Intra-grid*

This is a combination of various clusters. This form of grid computing is also known as a campus grid. It allows resources to be shared across various departments and units that function in terms of the same policies, without any need to address the security and policy management issues relating to global grids.

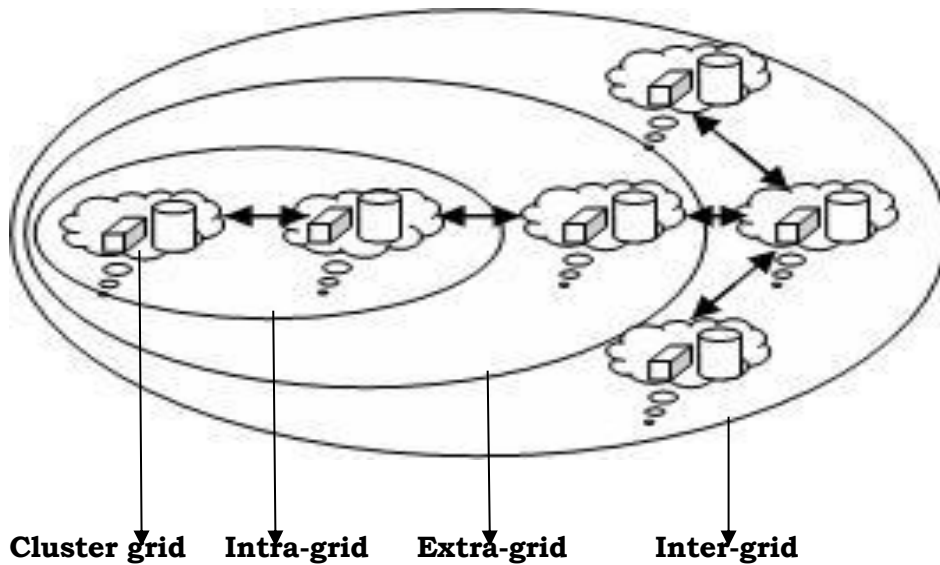


Figure 2: Grid classification based on topology

#### *Extra-grid*

This type of grid is also called a “partner grid” or an “extraprise grids”. It is a combination of two or more intra-grid with various security domains. This form of grid is geographically distributed among various establishments, companies and organizations. Virtual private networks (VPN) are used for implementing this type of grid to make resources easily available to users (Alfawair, Aldabbas, Bartels, & Zedan, 2007).

#### *Inter-grid*

This type of grid is the final phase of grid evolution. It is otherwise known as a global grid. It is a combination of both the intra-grids and cluster grids joined together by the Internet. This form of grid is popularly used in academia.

### **SECURITY OF THE GRID**

Security simply refers to measures put in place to ensure the safety, continuity, reliability, confidentiality and integrity of resources on a grid system. It can also be regarded as a means of safeguarding information and resources against illegal access, disclosure and alteration. Security is a very important concept that has been a serious research area in computing because of its effect on all the resources that are shared on a network.

*Security requirements in a grid environment*

In order to shield the resources on the grid from being attacked and from unlawful visitation, reliable privacy, data integrity, data confidentiality, non-repudiation, availability, authorization and authentication must be assured and provided ( Lu, Cao, Chai, & Liang, 2008).

The security requirements as given in Al-Bayatt, Zedan and Siewe (2009) were defined by the International Organization for Standardization (ISO) such as the International Telecommunication Union (ITU) e.g. ITU-T Provisions X.805 and X.800.

**Sensitive information**

This is information that should be accessed, viewed and updated by the authorized user alone. It is considered to be sensitive simply because it is not meant for public consumption and is for the use of a specified user or a group of users who have authority to do so. Any attempt to make it vulnerable to an outsider or an unauthorized fellow for manipulation and alteration renders the information's sensitivity and confidentiality worthless.

**Authorization**

For any organization to allow its resources to be jointly shared among all the parties involved, there is need for authorization in terms of who should have access to particular resources and who should not. It also allows permission for use to be given only to authorized nodes on the network (Al-Bayatt, Zedan, & Siewe, 2009). Globus Toolkit Gridmap files, the Community Authorization Service (CAS) and Virtual Organization Membership Service (VOMS) are authorization measures usually adopted in grid computing (Chadwick, 2005).

*Authentication and access control*

Impersonation has been identified as a major threat in a grid environment. Authentication is important to purposely prevent illegal access to resources ( Lu, Cao, Chai, & Liang, 2008). The main purpose of authentication is to confirm that the user is indeed the person whom he/she claims to be.

In both the shared and personal computer system, authentication is usually carried out with the use of a password and username. It has been established that when a password is used to log into the system, the authenticity of a user is usually fully guaranteed. However a password can be stolen, hence the information on the system can be vulnerable.

Digital certificates, verified by a certificate authority, are taken as the best way to ensure authentication on the Internet (Lu, Cao, Chai, & Liang, 2008).

### *Data confidentiality*

The purpose of data confidentiality is to protect data from being divulged to the wrong or an unintended party (Shen, Yan, Qiang, Wu, & Zhang, 2006).

Two processes can be used to achieve data confidentiality: data encryption and data decryption. Also, two main types of cryptography (Menezes, Vanstone, & Van Oorschot, 1997) can be used to provide data confidentiality (MSDN, 2005), i.e. symmetric and asymmetric cryptography.

### *Symmetric cryptography.*

In this type of cryptography, both the sender and the recipient use a common key to carry out encryption and decryption.

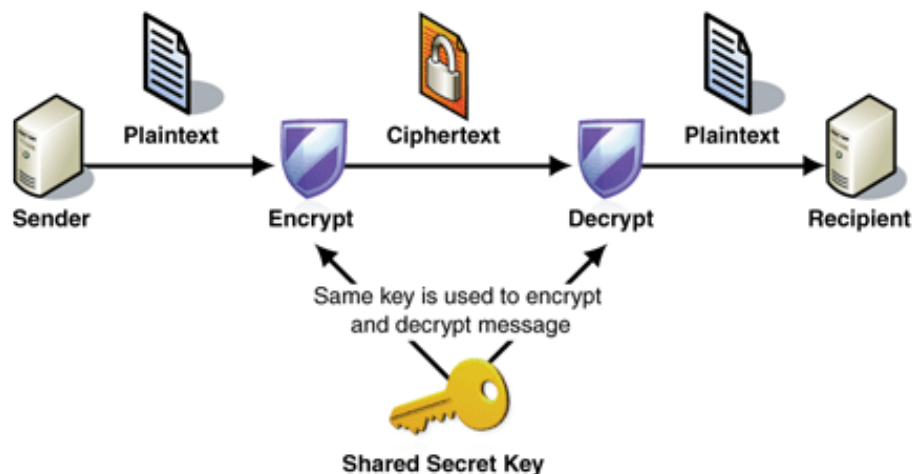


Figure 3: *The process of symmetric encryption*

As illustrated in Figure 3, symmetric encryption involves the following stages:

1. The cipher text message is created by the sender through the encryption of a plain text message with the assistance of a symmetric encryption algorithm and a shared key.
2. The cipher text message is sent to the recipient by the sender.
3. The cipher text message is decrypted back into a plaintext by the recipient.

#### *Asymmetric cryptography*

With asymmetric cryptography, which is also called public key cryptography, different keys are used by the sender and recipient for encryption and decryption, respectively (Cai, Liu, Guo, Zhang, & Geng, 2009). The sender encrypts data with one key, and the recipient uses a different key to decrypt cipher text (see Figure 4).

As illustrated in Figure 4, asymmetric encryption involves the following steps:

1. The cipher text message is created by the sender, who encrypts the plaintext message with the aid of an encryption algorithm and the recipient's public key.
2. The cipher text message is sent from the sender to the recipient.
3. The cipher text message is decrypted back to plaintext with the aid of a private key that tallies with the public key that was used to encrypt the message.

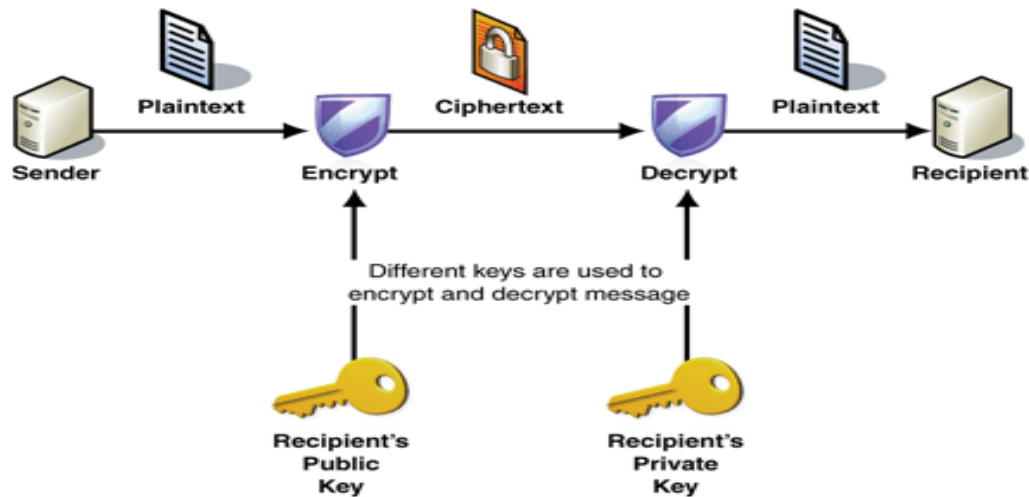


Figure 4: *Asymmetric encryption*

#### *Data integrity*

In the banking industry, the military and the aviation industry, data modification by an unauthorised person may have serious consequences (Akimana & Markowitch, 2006). With data integrity (Gilbert, Abraham, & Paprzycki, 2004), data in a grid environment is guaranteed to be removed, updated, modified, deleted, edited and transmitted only by an authorised person (Khider, Osman, & Sherkat, 2010).

#### *Non-repudiation*

Since transactions often take place on the Internet, repudiation prevents the parties involved in a transaction from denying that a particular transaction has taken place when it has indeed been carried out (Zhou, Deng, & Bao, 1999).

Non-repudiation therefore ensures that both the receiver and the sender cannot deny that a message has been sent or received. This security measure can assist in identifying and isolating any node on the grid that has become compromised (Onieva & Zhou, 2008).

This security measure is usually achieved through the use of digital signatures and certificates. Timestamps that contain the date and time can equally be used. Finally, a message transfer agent can be used to create and provide a digital receipt to establish that a message was sent and received (Anderson, 2008).



### *Privacy*

The main purpose of privacy is to ensure that information being shared on the grid system is protected. Every grid user wants his/her sensitive information to be completely secured from misuse and abuse ( Kaga & Abelson, 1990). The definition of privacy is as follows:

*"Privacy is the right of individuals to determine for themselves when, how, and to what extent information about them is communicated to others" (Ashley & Karjoth, 2003)*

### *Availability of resources*

Despite possible security attacks, data must be readily available across the network to satisfy the demand of grid computing users at any point in time (Chon, Enokido, Wietrzsk, & Takizawa, 2004). Data availability means that data is available at all times. In a grid environment, data availability is usually achieved through redundancy, which has to do with how data is stored and how it can be reached. Also, essential and adequate services must be made available by a node at any time (Berman, Fox, Hey, & Hey, 2005).

### **Authentication**

Authentication simply means the process by which an intended user of a computer system is verified thoroughly to confirm if there is a legal backing and allowance for his/her usage of or access to the system (Laccetti & Schmid, 2007 ). Some of the approaches that are generally adopted for carrying out authentication include a username and password (Federico, 2009). They can also include other methods of demonstrating identity such as voice recognition and fingerprints (Dewan, Grundin, & Horvitz, 2007).

### *Attacks on a grid system*

It has been recognized by researchers that all forms of distributed systems are vulnerable to attacks and not completely secured. Some of the security measures used are inflexible and not scalable (Laccetti & Schmid, 2007 ).

Security attacks on a grid system can be classified as follows:

- Passive attacks;
- Active attacks; and
- Dictionary attacks.

#### *Passive attacks*

Passive security attacks ranges from secret monitoring of transmissions, such as electronic mail messages, file exchange on any distributed system, client-server transmission or eavesdropping (Al-Bayatt, Zedan, & Siewe, 2009). Passive attacks involve exposing and releasing of message content and a thorough analysis of traffic:

#### *Exposing and releasing of message contents*

This form of attack is explained in Figure 5. It is obvious that electronic mail message, telephone conversations and a file being transferred may contain sensitive and confidential information (Stallings, 2002). To guard against the vulnerability of the information, it is necessary to prevent someone who is not authorised to do so from learning and understanding the contents of the transmission (Welch & Lathrop, 2003).

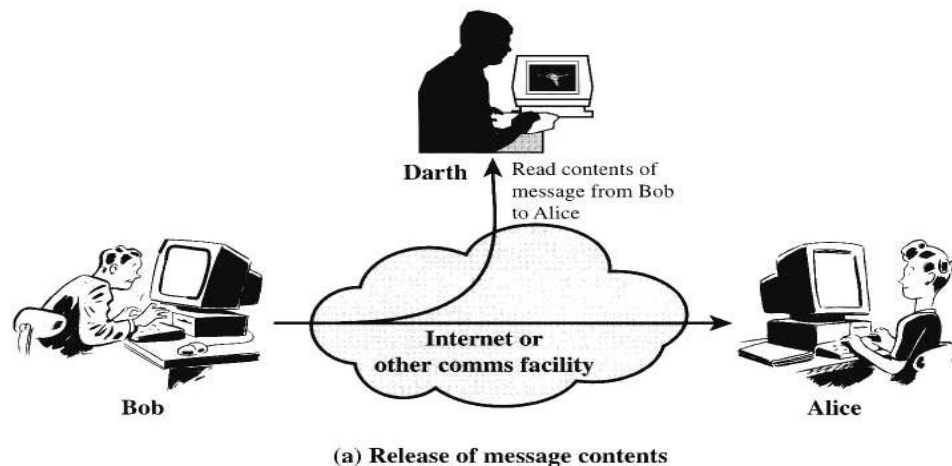


Figure 5: Release of message contents (passive attack)

#### *Traffic analysis*

This is a form of security attack whereby messages are intercepted and examined in order to acquire information from patterns in the communication process. This form of attack can

be carried out even when the message is encrypted and is difficult to decrypt (Chon, Enokido, Wietrzsk, & Takizawa, 2004). As a general principle, the higher the number of messages intercepted and examined the more the deduction and inference from the traffic. Traffic analysis can be carried out by military intelligence and it is a serious concern in computer security (McClure, Scambray, & Kurtz, 2003). Supposing a message is masked (encrypted), an adversary might still gain valuable information from the message by determining the location and identity of the communicating host and the frequency with which messages are exchanged and their length (Khider, Osman, & Sherkat, 2010). It has been established that passive attacks are extremely difficult to detect since they do not involve the alteration of the data in the messages.

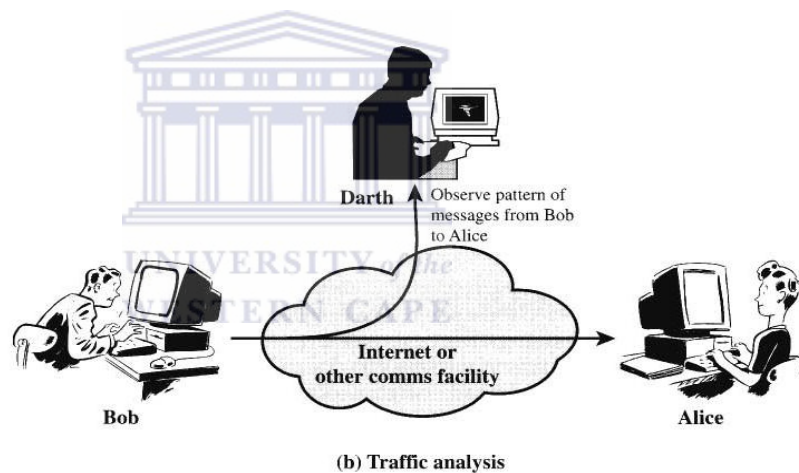


Figure 6: Traffic analysis

#### Active attacks

This type of attack attempts to change, remove, or destroy data being transferred from one system to another on a grid network. Some of the types of active attacks are denial of service (DoS), masquerade, replay and modification. An active attack can be internal or external (Welch & Lathrop, 2003). It can be prevented with the aid of common and popular security mechanisms like firewalls and encryption techniques. The section below gives brief explanations of some of the examples of active attacks.

*Denial of service (DoS)*

This type of active security attack halts the system on a grid (IASTED, 2004) so that it fails to respond to users' demand. To accomplish this objective, attackers will send a very large quantity of data at the access point (Mohteshim, 2005) so that it makes it extremely difficult to respond appropriately. DoS deliberately aims at preventing legitimate grid users from accessing the resources on the grid.

*Masquerade or impersonation*

In this case both the attitude and behaviour of an authorized grid user are mimicked and copied by the attackers. With this approach, attackers can easily modify and change information to the detriment of a legal user. Man-in-the middle attack is the commonest form of impersonation.

*Disclosure*

Sensitive information across the grid might be disclosed by a compromised machine on the grid to another machine that is not authorized to have access to such information. Also, classified information might be divulged by a user that is not recognized to access it, thus compromising the secrecy of the information (Al-Bayatt, Zedan, & Siewe, 2009).

*Unauthorized access*

This type of security attack occurs when a person who is permitted to access particular information gains access to it and interfere with it in some way as if he were the rightful owner of the information. When this happens the sensitivity of the information will be lost and it can easily be manipulated without challenge (Imine, Cherif, & Rusinowitch, 2009). This may happen through unpatched software or other known vulnerabilities.

*Replay attack*

Replay attack can also be likened to the man-in-the-middle form of attack. This form of security attack allows the data packets to be intercepted and replayed (resent) to the server. For example, if a client sends an encrypted password and a username to a server to gain access to any grid information and a hacker is able to use a monitoring software to intercept

such a message and replay/resend it, such a hacker will have the same right of access as the rightful owner the message (Jeong, Yu, Shin, Shin, Moon, & Lee, 2006). In fact, the hacker may even attempt to alter the password and thereby deny the authorized user access to the information. This type of attack can be avoided with the aid of digital signatures.

#### *Dictionary attack*

This method is used by attackers to break a security measure put in place for a system or a server by trying all possible passwords with the intention of gaining access. The attackers begin by using the most common way people create a password, e.g. by using their names, children's names, date of birth and place of work. The attack is carried out with the assistance of software (Azeez & Osunade, 2009). This form of attack allows words to be sorted by usage frequency and begins with the most likely one.

Dictionary attacks allow spammers to randomly send mail across to various addresses using a combination of some popular domain names with the intention of sending information across to a large number of e-mail users (Pinkas & Sander, 2008). For example a dictionary attack may begin with the following e-mail addresses john@uwc.ac.za, john1@uwc.ac.za, john2@uwc.ac.za, john3@uwc.ac.za. This may continue until all possible combinations of both the numbers and variable have been tried.

Account locking and delayed response are two major countermeasures against dictionary attacks. Account locking allows the accounts to be locked completely to the attackers after they have made a few unsuccessful efforts to gain access (Houmb., Georg, France, & Matheson, 2004). In delayed response, a server gives a slow YES/NO response when a login-name/password is supplied. This prevents attackers from rapidly checking a large number of passwords (Pinkas & Sander, 2008).

#### **Architecture**

This is a structural representation and design framework of how a model works. Architecture in this context implies a corresponding flow of information with regard to the way the model explicitly spells it out. It varies from one model to another.

### **Prototype**

This is a framework to accurately simulate the characteristic features of a planned design. The main reason for having a prototype is to attain an agreement level between the user and the designer/developer to ensure whether the design conforms to specifications and, if it does not, to make amendment before a final decision is taken on such a design (McLean, 2008). A certain level of expertise is very crucial to satisfactorily employ prototyping as a design checking tool. Prototyping testing can be used to bring down the level of risk and hazard a design could generate.

### **Computational grid**

A computational grid is a form of grid system that allows the connection of various hardware and software resources from various organizations so as to provide sharable resources across a multiple administrative domain and to cater for users' needs in a reliable and secure manner (Brown, 1999). In this type of grid system, resources are usually aggregated in order to act as a unified processing resource (Machiraju, Sahai, & Moorsel, 2003).



### **Security model**

This can be regarded as a system of stating and implementing security policies. A security model (Goguen & Meseguer, 2009) may be based on formal model access rights or a model of distributed computing (Geogiev & Georgiev, 2001). This is a statement of the requirements crucial to adequately support and implement a certain security policy. If there is a security policy that requires authentication and authorization before permitting the accessing of information on the network (Buttler, et al., 2000), the security model might schedule an access control matrix that should be designed so that it supports and fulfils the requirements of the security policy (Pinkas & Sander, 2008). It might be possible for a security policy to state that no one from a lower security level should access or view particular information at a higher security level. The security model in this regard will state adequate rules and guidelines that are expected to be put in place for this security policy to be implemented (Ashley & Karjoth, 2003).

**Triple-domain grid-based environment**

This is an architectural model designed to cater for resource sharing among three levels of domains. Each of the domains is characterised by the local security monitoring unit which is responsible for the authorisation and authentication of users and a central security unit that allow all users on the grid to be given final authorization in order to have access to resources on the network.

In this dissertation, scalability, interoperability and efficient access control are simultaneously addressed because of low attention to addressing these security challenges and their importance in grid computing.

**CONCLUSION**

Chapter 2 contains a literature review in terms of the research questions posed in Chapter 1. Explanations of various categories of grid systems and their categorisations based on their topology were presented. This chapter also explains the various security requirements of a grid-based environment in some detail, while the various kinds of security attacks are also discussed. In the next chapter (Chapter 3), the research design and methodology are presented, which will in turn provide insight into analysis of results presented in Chapter 4.

## *Chapter 3*

### RESEARCH DESIGN AND METHODOLOGY

#### **INTRODUCTION**

In the previous chapter, literature related to the research was presented. The focus areas included security, access control, scalability and interoperability. In the present chapter the research design and methodology that were applied in the study will be discussed. The chapter is divided into two main sections: the first covers the research design while the second discusses the methodology that was used. The results obtained in this work are divided into three categories: results on scalability, results on interoperability and results on access control.

The results of scalability were achieved through the use of a simulator (G3S) by measuring the entities (number of accessed resources against time, average turnaround time against number of grid requesters and average turnaround time against number of grid requesters) and by evaluating the flexibility of the 3DGBE architecture with one of the traditional access control models, mandatory access control (MAC).

#### **RESEARCH DESIGN**

Crotty (1998) defined the research process in terms of four elements: the epistemological stance, the theoretical perspective, the methodology and methods deployed (see Figure 7).

##### **Epistemology**

Epistemologists acknowledge four main channels of knowledge. They are intuitive knowledge, authoritative knowledge, logical knowledge and empirical knowledge. According to Cline (2011),



*“Epistemology is the investigation into the grounds and nature of knowledge itself. The study of epistemology focuses on our means for acquiring knowledge and how we can differentiate between truth and falsehood”.*

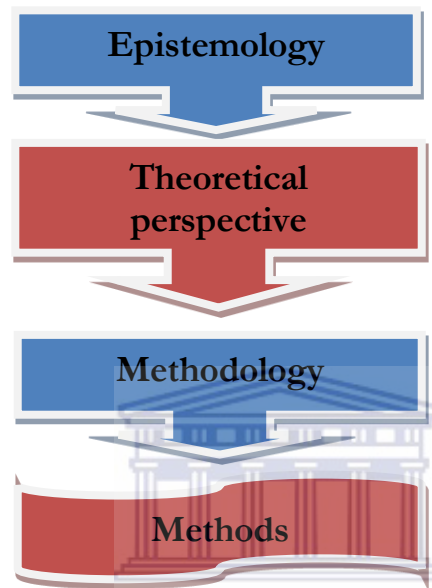


Figure 7: Four elements of the research process (Crotty, 1998, p. 37)

The research was viewed from the epistemological stance of objectivism. Objectivism rejects the notion that a group of people or individuals establish their own reality without verification (Dewan, Grundin, & Horvitz, 2007). With objectivism as the dominant epistemology, the assumption is that user evaluation and judgmental results can be evaluated, verified and quantified (Crotty, 1998).

### **Theoretical perspective**

The other aspect of the research design is the use of theoretical perspective:

*“A theoretical perspective is a non-explanatory general framework. It is meant to define a point of view within a discipline, which may include basic assumptions that draw attention to aspects of a phenomenon. It will then generate questions about it” (Darke, Shanks, & Broadben, 1998).*

Positivism is a theoretical perspective that allows for systematic, practical and empirical evaluation of a natural occurrence that is based on scientific theory and hypotheses about interactions among such occurrence (McCroskey & Richmond, 1990).

In the case of this research, positivism is appropriate, mainly because the research required a scientific or qualitative appraisal (Bogdan & Taylor, 1998) of the model and the algorithm that was developed.

### **Methodology**

Methodology is a strategy or action plan to choose appropriate research methods. In this study, soft-system methodology (SSM) was used to manage the research process.

Soft system methodology (SSM) was conceived by Peter Checkland of Lancaster University (Checkland, 2000) as a methodology suitable for research problems with no obvious or clearly defined solution, i.e. it is a methodology suitable for solving messy and intricate problems.

The basic shape of SSM (shown in Figure 8) is a cyclical process. The A in this diagram depicts a real-world situation of concern. This first phase of the SSM cycle helps to explore and define the problem situation. The second phase of the SSM cycle (B in Figure 8) depicts relevant systems of purposeful activities. These “purposeful activities” should address the problem situation of “How should information on a grid be secured without compromising the accessibility and the availability of resources?”.

*The comparison of models with perceived real-world situation* (see C in Figure 8) is considered as the powerhouse of this methodology. At this level the model is compared with reality and insights are drawn from the comparison. Metrics such as turnaround time, grid nodes, users, time, throughput and grid service requester are used for this comparison.

*The action needed to improve the situation* (see D in Figure 8): is the phase of SSM that determines what action is needed to improve the problem situation.

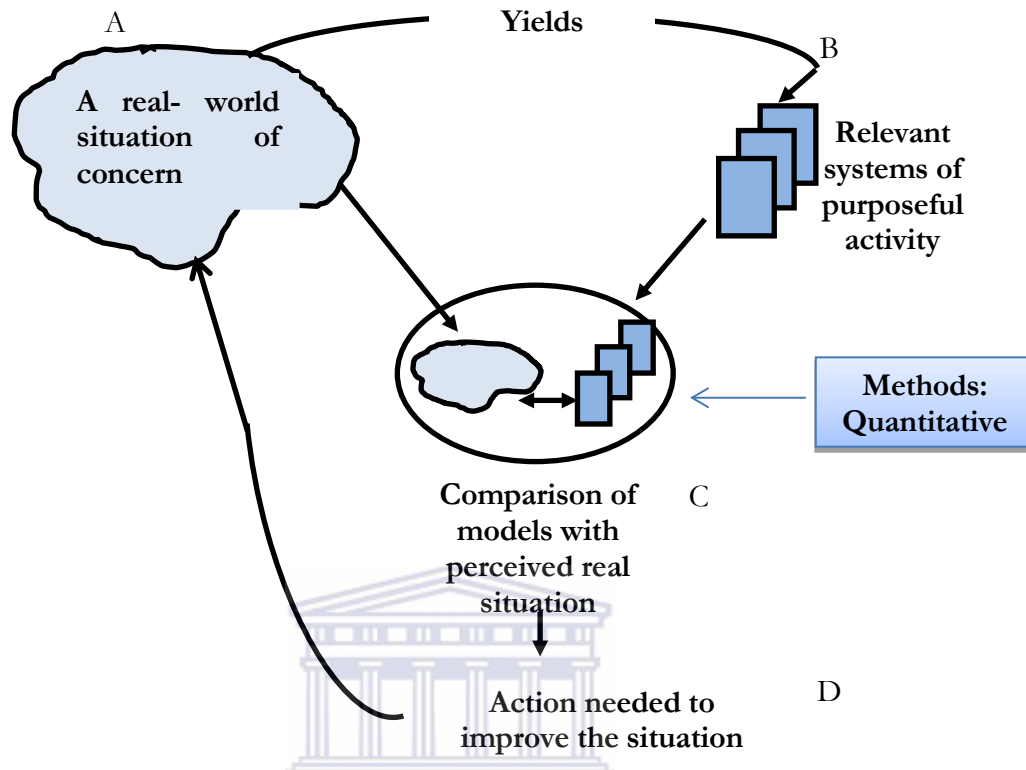


Figure 8: The basic shape of SSM (Checkland, 2000, p. 23)

For this research, the following approaches were followed (see Figure 9).

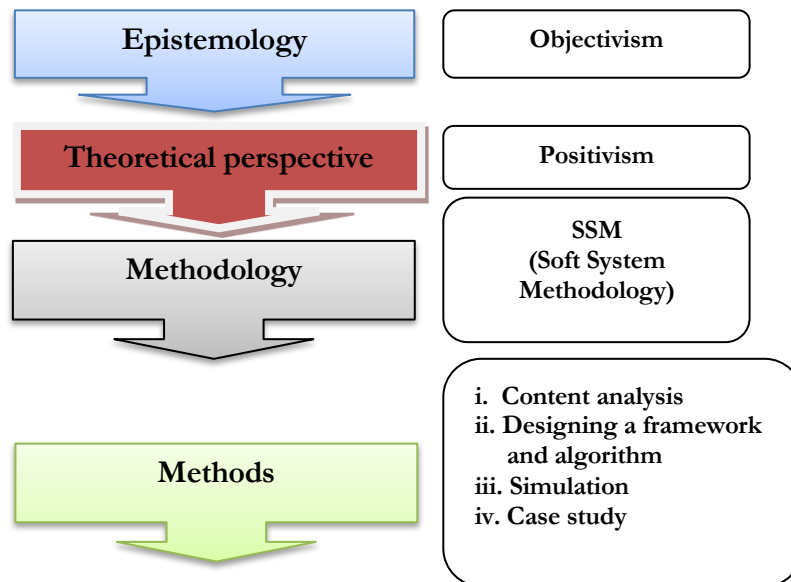


Figure 9: The four basic research approaches used in this thesis

## Methods

### *Content analysis*

According to Slarkin, (2010), content analysis explains the act of reviewing the existing documentation of related research areas so as to retrieve and extract items of information that are useful to the current research and project, hence it should be regarded as an important project requirement.

The analysis was carried out using the interpretive approach. According to Mather (2010), the interpretive approach allows a reviewer to read, digest and interpret a given article or document so as to draw connections between these documents and the research area currently being studied (Cassell & Symon, 1994). Content analysis was carried out to determine the strengths and weaknesses of five security models in order to formally establish their scalability, interoperability and efficient access control

### *Architectural framework and algorithm design*

To find a solution to the identified challenges, an architectural framework was developed. A two –stage architecture based on the RBAC model was developed. Stage 1 allows grid users to interact with the local security monitoring unit (LSMU) and central security monitoring unit (CSMU) for purposes of authorisation and authentication. Stage 2 of the architecture makes use of a policy information point (PIP), policy enforcement point (PEP) and policy decision point (PDP), all of which are extensible application markup language (XAML) protocols to verify and confirm the grid user based on his/her roles.

A framework is a pictorial and structural presentation and representation of a proposed architecture.

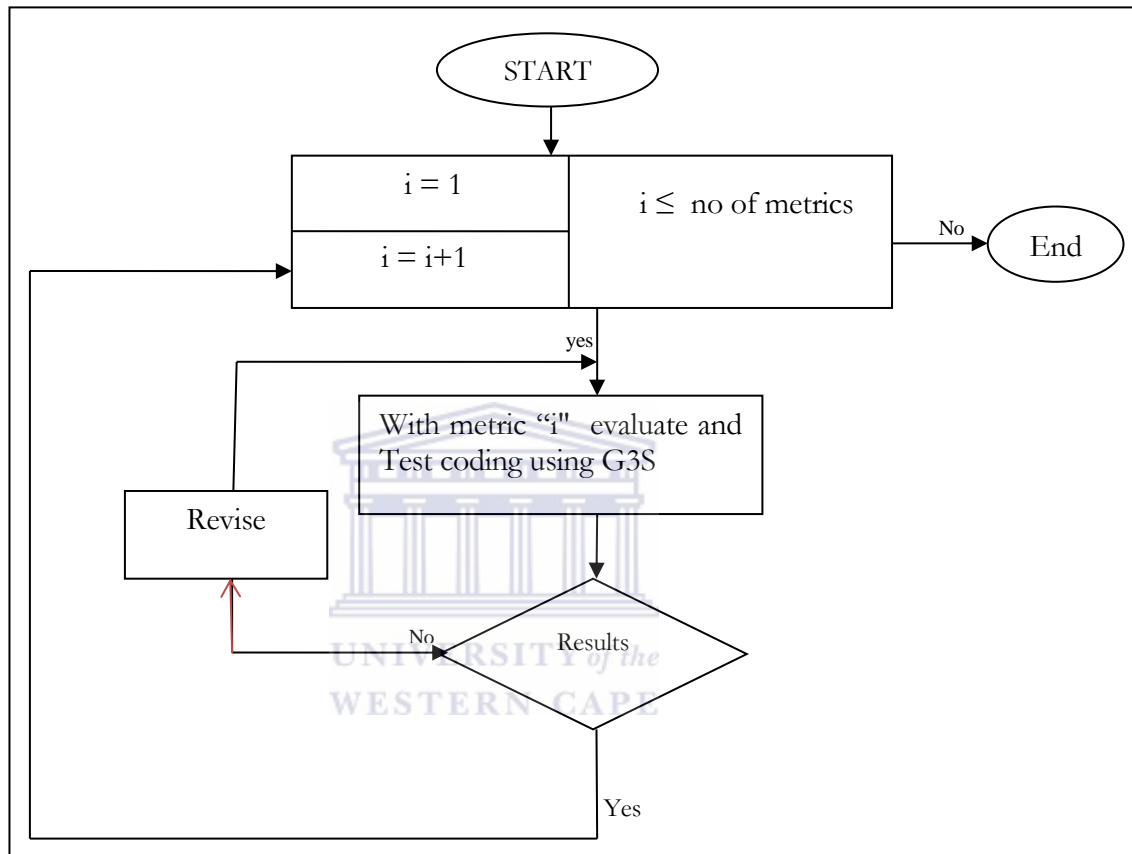


Figure 10: Simulation cycle using G3S

#### *Simulation and experimentation*

A grid security services simulator (G3S) was used for simulation and evaluation of the framework. G3S was used because it provides adequate support and a suitable mechanism for simulating security functionalities (Naqvi & Riguidel, 2005). The simulation cycle is depicted in Figure 10.

The fact that the efforts to obtain reliable and consistent results are good sign of the authenticity of the achieved results and, hence, several testing were conducted. The objective of the designed framework is to efficiently handle the identified problems during

content analysis. Simulation and experimentation were used to evaluate the efficiency of the frameworks and algorithms (Figure 10). The results obtained are presented in phases based on the SSM methodology adopted.

### APPLICATION OF SSM TO THE RESEARCH PROBLEM

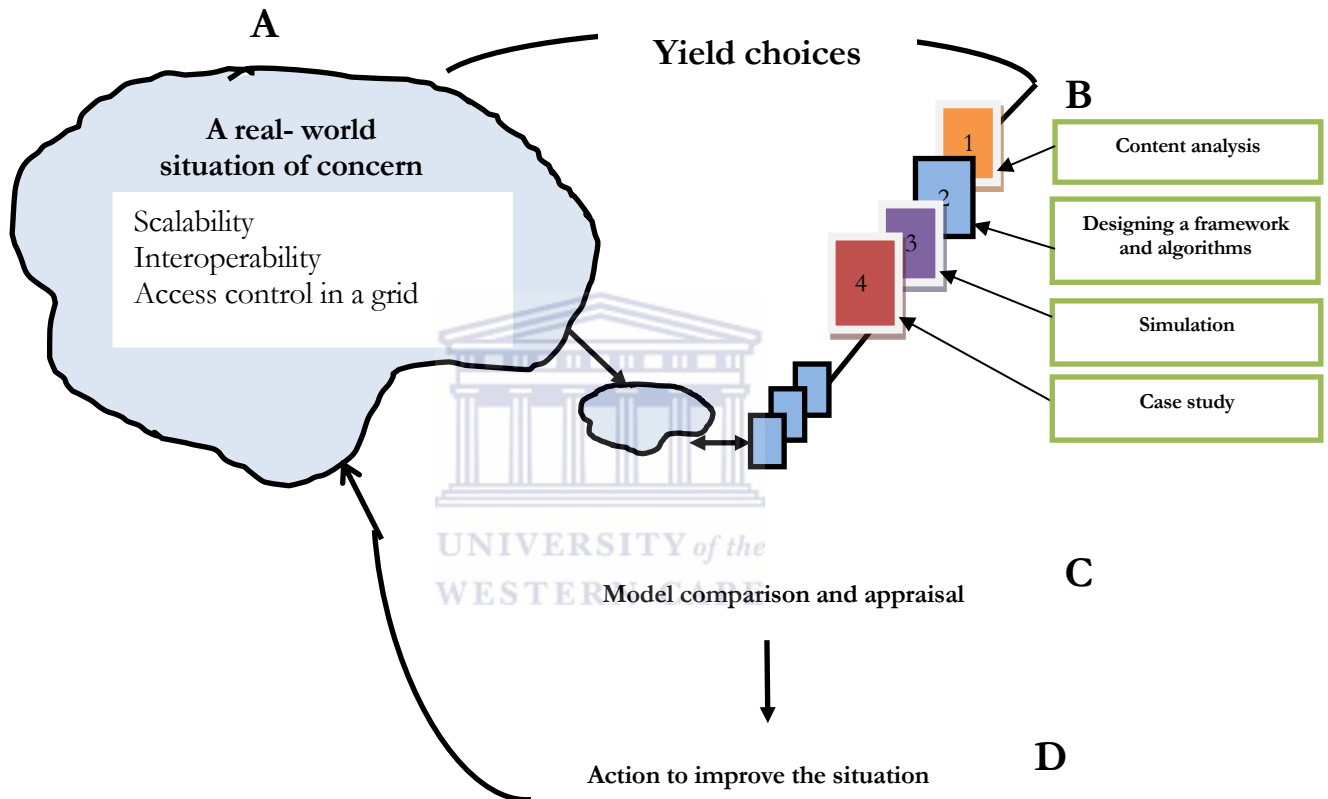


Figure 11: Application of the SSM, adapted from (Checkland, 2000, p. 23)

The implementation stages of the SSM methodology (as shown in Figure 11) was divided into four phases. The first phase dealt with the identification of challenges using content analysis; the second dealt with designing the architectural framework and algorithms and the third and fourth stages dealt with simulation and experimentation, as well as a case study used for implementing access control. Each of the stages was repeated several times in order to obtain consistent and reliable results.

### *Grid security services simulator (G3S)*

The evaluation of a complex grid architecture like the one being considered (3DGBE) might not be feasibly implemented and carried out in a real grid environment due to the flexibility and dynamism of its nature.

Grid testbeds are very expensive and time consuming because they require various policies to be specified for entities ( Sulistio, 2008, p. 24). As a result, it is easier to make use of simulator to study and evaluate the complex and intricate scenario.

The reason for the use of simulators as an alternative to the expensive and tie consuming test beds is valid and the efforts have been put to acquire closer to real results. Grid security services simulator (G3S) has been used.

Table 1 provides information on the available grid simulators and their corresponding applications.

Table 1: *Grid simulators with their respective applications*

No.	Simulator	Application on grid	References
1	<b>SimGrid</b>	For compiling time and scheduling algorithms	Legrand et. al, 2003
2	<b>GridSim</b>	Scheduling and pricing	Buyya & Murshed, 2002
3	<b>GangSim</b>	Synchronous and asynchronous workloads	Dumitescu & Foster, 2005
4	<b>OptorSim</b>	Stability and transient behavior of replication	Cameron , 2004
5	<b>G3S</b>	Security, access control, and scalability	Syed & Michel, 2005

Except for G3S none of the simulators is capable of handling any of the security challenges such as interoperability, policy level, authentication, usability, confidentiality, trust, scalability

and authorization like G3S. This is because; G3S is specifically designed to handle security-related matter in a grid environment.

It was developed by Syed Naqvi and Michel Riguidel of Computer Sciences and Networks Department of Ecole Nationale Supérieur des Télécommunications (ENST), Paris, France. The simulator was released for the experimentation by the developers based on the special request by the student and motivation from the supervisor in July, 2010.

G3S 2.0 is a simulator that allows simulating security functionality in a grid based environment. It simulates functionalities such as authorization based on the RBAC security model, authentication, and confidentiality based on the Bell LaPadula model. The simulator has a graphic user interface (GUI) that allows it to register grid users, creates various homogeneous groups, and simulates cases of interaction and communications between various entities to reveal and depict how the virtualization of security functionality can exist. Aside from being able to simulate the above mentioned functionalities, G3S also has some simulation functions like distribution of trust, attack patterns, secure exchange of documents and the security policy of the grid.

As shown in Figure 12, G3S has various buttons that are available for adding new users and new nodes, creating VOs and attack patterns, establishing secure exchange of documents etc. This ensures scalability/extendability.



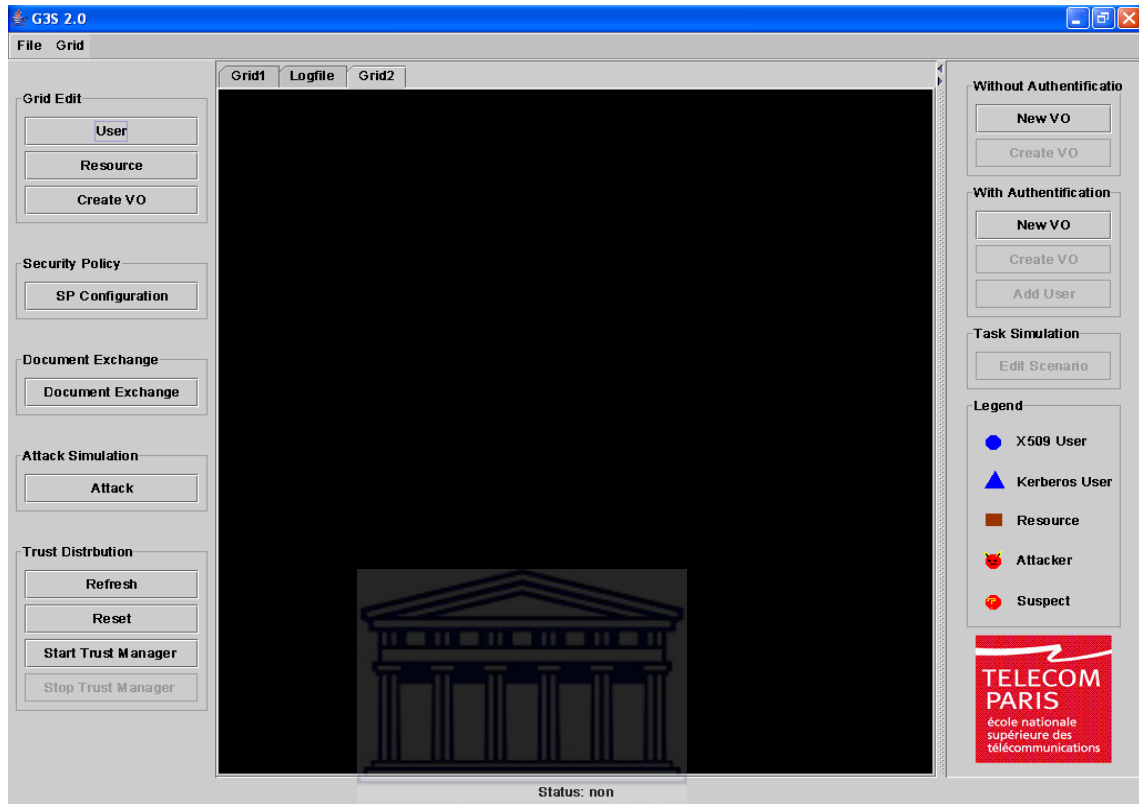


Figure 12: Graphics User Interface (GUI) of G3S

The simulator allows various security policies to be configured (see Figure 13). Both the security services and policies are easy to create with the aid of this GUI. A VO is selected and the configurations are set and saved accordingly.

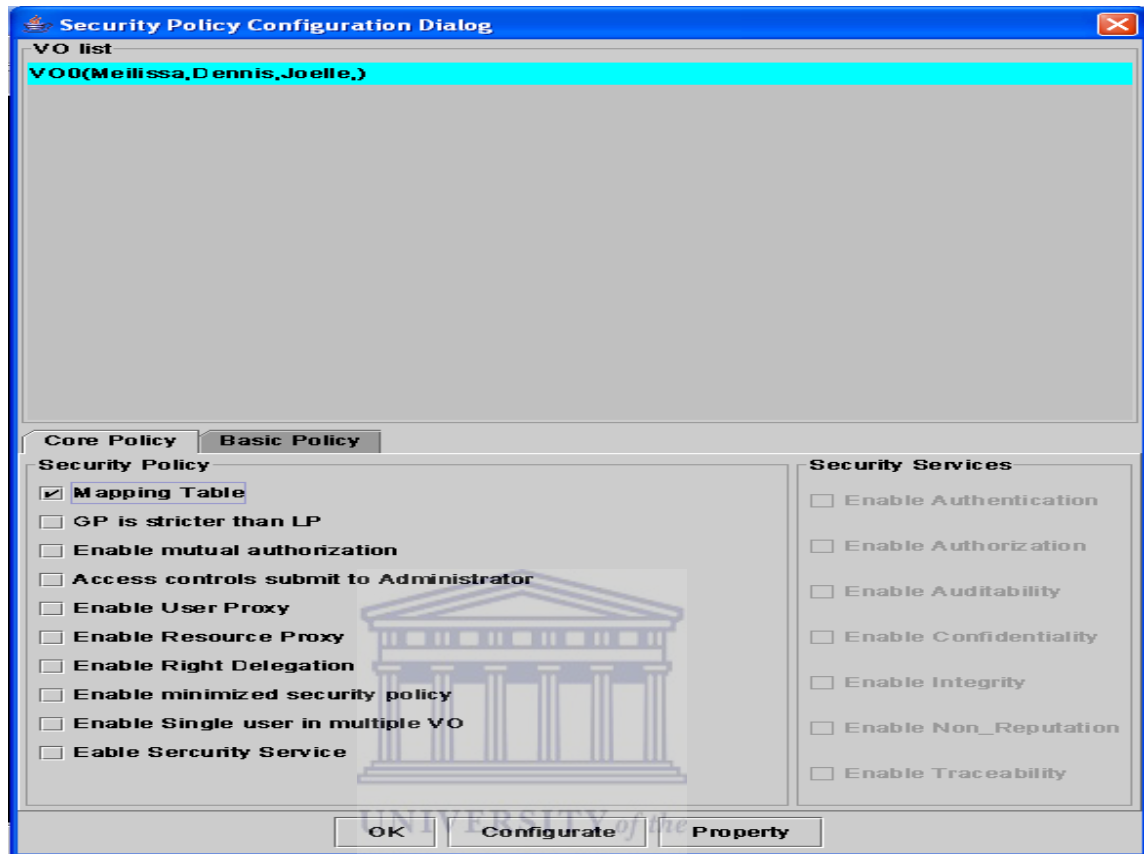


Figure 13: Configuring security services and policies

With G3S, new computing resources can be added and upgraded while new groups and grid users can be dynamically included into the network. Also, new VOs can be established and various security services and policies can be dynamically configured.

#### *Metrics used in the simulation*

It is worth mentioning that each of the challenges addressed in this thesis were evaluated in terms of different criteria. Scalability was evaluated using the following metrics: number of grid nodes, time (sec; see below), number of accessed resources, average turnaround time and number of service requesters, among, others while the method used was simulation.

Interoperability, on the other hand was addressed and evaluated by using middleware, the operating system, a federated database and authorisation between LSMU and CSMU, while case study and simulation were the methods used.

Roles, services, permissions and hierarchy were the metrics used for evaluating access control using a health case study as the method.

Metrics used in 3DGBE are referred to as the analytical measurement determined to measure, calculate and quantify the state of the architectures developed in this thesis. Apart from the definitions previously provided, below are the explanations of other metrics used for evaluation.

*Number of grid nodes:* This is the number of network junctions or network connection points. Every terminal, computer, hub and switch is a node. As used in the simulation, the number of grid nodes represents the number of network joints simulated for each of the domains.

*Time (sec):* Time measured in seconds is the exact time taken to access a certain number of resources when a user makes a request in a simulated 3DGBE.

*Number of accessed resources:* This is the number of resources granted a user within a given period of time.

*Average turnaround time (sec):* This is the time taken between the period of request for accessing resources in a 3DGBE and the return of detail request output to the grid user.

*Service requesters (grid users):* These are the people who access a grid network with the sole intention of gaining access to resources.

#### *Definition of simulation parameters for a triple-domain simulated grid based environment*

Three domains were considered so as to create and establish reasonable and manageable grid network among the entities (users and resources) sharing resources. In order to evaluate the effectiveness of the simulation of the domains, the parameters defined below were taken into consideration.

**Definition 1:** Let  $DSR(A, B)$ ,  $DSR(A, a)$ , denote the direct security rate which is determined and evaluated when the CSMU finds and grants permission and access privilege to a user from Domain B to Domain A or from an entity  $a \in \text{Domain A}$  to Domain A depending on

from where the access is requested.  $DSR(A, B, C)$  denotes the DSR between the three designated domains.

**Definition 2:** Similarly let  $SR(A, B)$  or  $SR(A, a)$  denote the security rate for accesses from Domain B to Domain A or for an access from entity  $a \in \text{Domain A}$  to Domain A.  $SR(A, B, C)$  denotes the SR between the three designated domains.

**Definition 3:** Let  $\text{Assess}(a_1 \dots a_m)^m$  denote assessment for entities  $a_1 \dots a_m$  when  $a_1 \dots a_m$  terminates at time step  $m$ , and  $-1 \leq \text{Assess}(a_1 \dots a_m)^m \leq 1$  shows either rejection or satisfaction during the assessment of the entities involved. While '-1' indicates the rejection that will reduce the value of SR, '+1', indicates satisfaction, which will increase the value of SR.

**Definition 4:** Let  $DSR(a_1, \dots, a_m)$  stands for “*Direct Security Rate*” in a grid for entities  $a_1, \dots, a_m$ .

**Definition 5:** Let  $\text{Rep}(A, a)$  denote reputation and status of entity  $a$  in Domain A on a grid.

**Definition 6:** Let  $\text{Approv}(a_1 \dots a_m)^m$  stand for the approval in the service request for  $a_1 \dots a_m$  after  $m$  time steps.

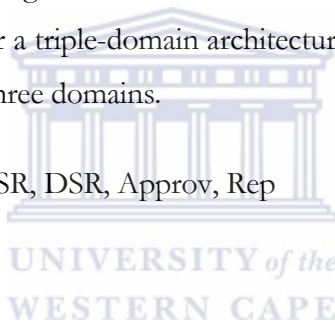
Determining or evaluating the security rate in a multi-domain grid-based environment is completely different from evaluating what is obtainable in a single-domain environment. The main reason for this is the interaction and relationship between the grid entities involved. Unlike in a single-domain environment, a multi-domain grid environment has more entities from one domain to another to interact with. Hence, to handle the complexities that arise from the user's accessibility to the resources of different domains, the SRs for the entities of each domain are useful for quick and accurate evaluation of the security within different domains. The approach adopted for determining the inter-domain security rate value is simple and provides the benefit of feedback that is flexible and dynamic in nature.  $\text{Rep}(C, a)$  yields status/repute of entity  $a$  to Domain C in a VO when considering that  $a$  is not an entity in Domain C. It is worth mentioning that A, B, and C

represent three different domains being considered while  $a_i$ ,  $b_i$  and  $c_i$  are entities in the three domains.

One of the main benefits of grid computing is to ensure a well-coordinated resource distribution and promote unified work between different domains in a flexible grid based environments. In order to safeguard and protect each GU's security and privilege, a dependable, secure and efficient access control is important.

Since one of the issues being addressed in this thesis is efficient access control, a new calculation approach based on security rate across domains is proposed and analysed by taking into consideration the grid entities across the three available domains in a 3DGBE. This approach is suitable for a triple-domain architecture and also adaptive for a category of architecture with less than three domains.

Parameters defined include SR, DSR, Approv, Rep



Hence,

$$SR(A, B, C) = \lambda_1 DSR(A, B, C) + \lambda_2 Rep(A, B, C) \dots\dots\dots(1)$$

Equation 1 is used to evaluate the SR in the three domains A, B, C

With Rep where the weight  $\lambda_1$  and  $\lambda_2$  are positive and  $\lambda_1 + \lambda_2 \leq 1$ .

$$DSR (A_i, A_j) = \frac{\sum_{a \in A_j} DSR(A_i, a)}{|A_j|} \dots\dots\dots(2)$$

Where  $a$  is an entity from the Domain A. Given two different domains  $A_i$  and  $A_j$  with  $i, j \in 2 [1 \dots n]$ , where  $i \neq j$ , and  $n$  is the number of domains.

Therefore,

$$DSR (A, C) = \frac{\sum_{c \in C} DSR(A, c)}{|C|} \dots\dots\dots(3)$$

When considering any domain, A, B or C, Equation 2 is generic and can therefore be used to compute DSR between them. The same is applicable to Equation 3 where Domains A and C were specifically considered.

For Domains  $A_i$  to  $A_j$  with  $i \neq j$ , the status of entities is determined as follows:

$$Rep (A_i, A_j) = \sum_{a \in A_j} \theta_a Approv (A_i, a) Rep (A_i, a) \dots \dots \dots (4)$$

where  $\theta_a > 0$  is the weight given to  $Approv(A, a)$  for  $a \in A$  and  $\sum_{a \in A} \theta_a = 1$ . Equation 3 implies that the  $Rep$  can be determined from any desired domain and can be extended to any number of domains.

$Rep$  is the evaluation of other grid entities from any unusual and unrelated source. The values of the weighted factors are associated with the requirements of local security policies in each domain.



$Approv (a_i, a_j)^k$ , implies the Approval degree of service for  $a_i$  to  $a_j$  after they finish  $K$  times services.

Hence,

$$Approv(a_i, a_j)^n = \lambda_1 * Approv (a_i, a_j)^{n-1} + \lambda_1 * V (a_i, a_j)^n \dots \dots \dots (5)$$

$$Rep (A_i, a) = \frac{\sum_{a \in A_i} Approv(a_i, a_j)}{|A_j|} * Rep (A, a_i) \dots \dots \dots (6)$$

$Rep (A_i, A_j)$  is the reputation of domain  $A_j$  to  $A_i$ . This approach introduces a weight  $\Theta_a$ , and it corresponds with entity  $a$  in domain  $A$ . An entity with high secure rate is more likely to be given access permission than entity with lower security rate.

Hence,  $\Theta_a$  is proportional to  $SR (A_i, a)$ . A GU entity's action in a domain defines and influences its access ability to any other domain's resource, which is a clear and implicit restriction to user.

*Case study: federated database*

It should be recalled that interoperability was evaluated by using middleware, an operating system, a federated database and authorisation between the LSMU and CSMU.

This section explains briefly how a case study was used to achieve interoperability using a federated database.

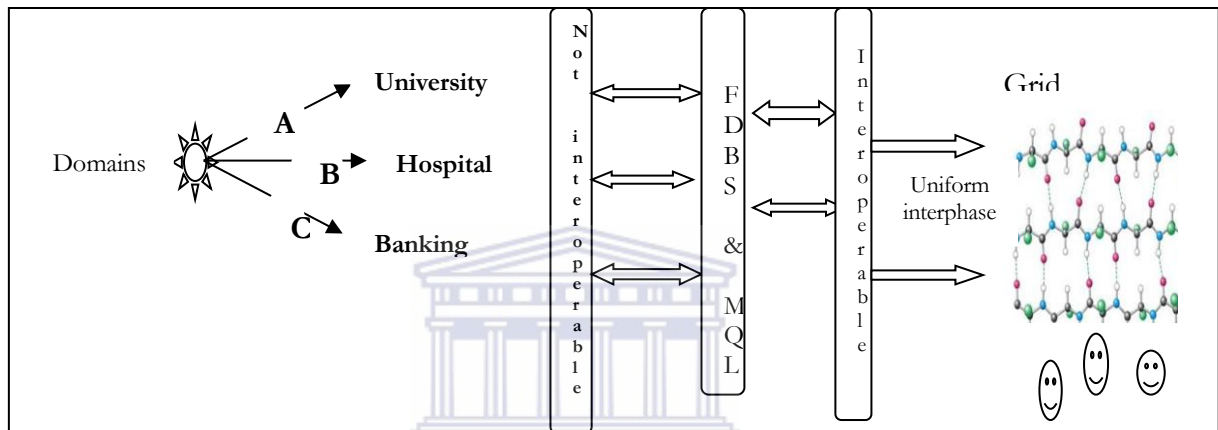


Figure 14: Schematic representation of 3DGBE interoperability operation flow

Major challenges concerning getting information from a multi-domain grid environment is the variation of data formats and the semantics adopted by the database (Vermeer & Apers, 1996). As a result, the need for data interoperability for ensuring uniform interpretation and analysis of the retrieved information from different domains is important. All the domains should submit to a uniform information reference model. As shown in Figure 14, the need to aggregate information from different domains is due to various applications being run on them.

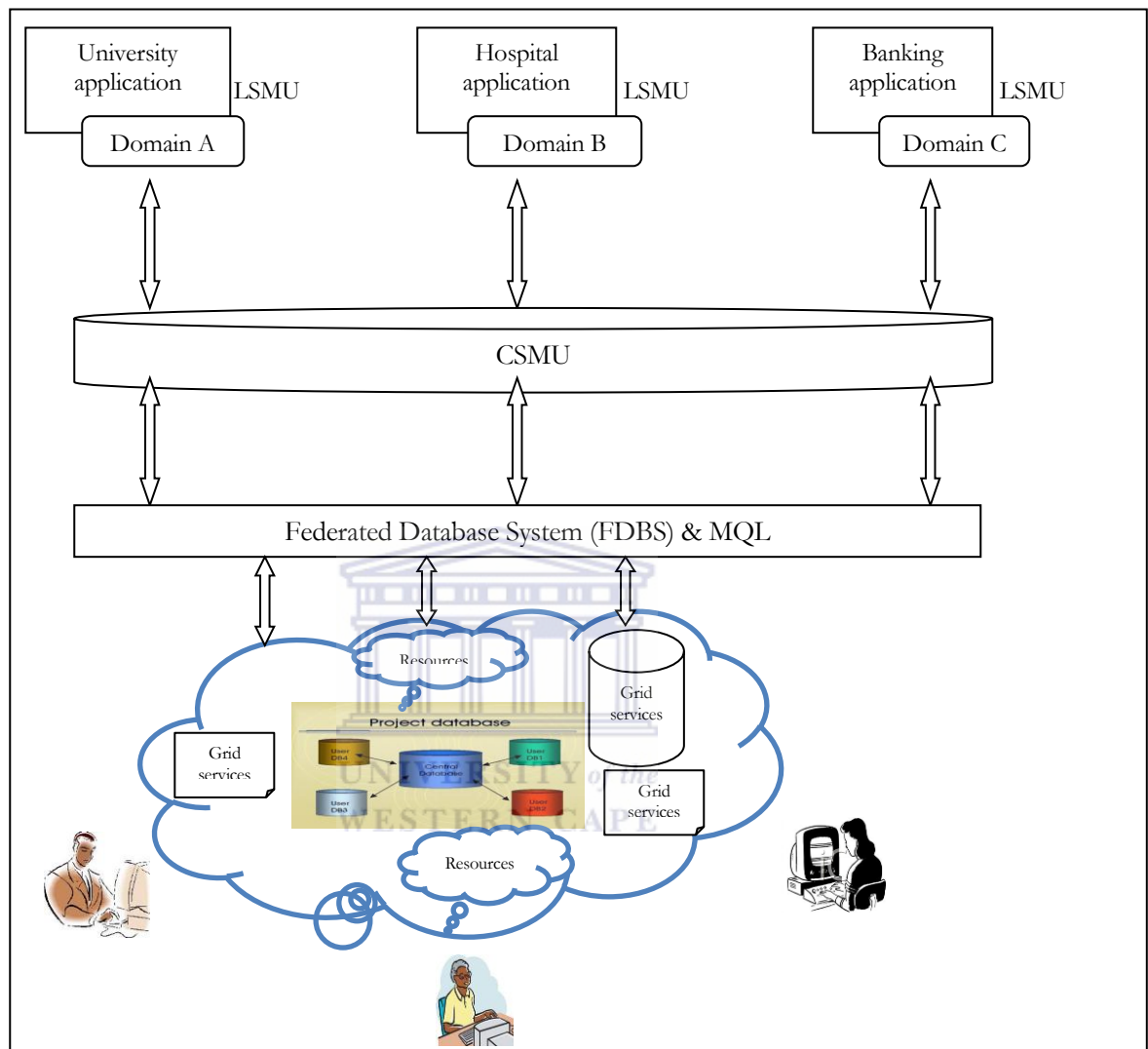


Figure 15: Gridification of the 3DGBE database architecture for interoperability

The developed framework consists of a three-domain application-sharing environment for University, Hospital and Banking. (see Figure 15) A set of classes, attributes and class-relationships between the three domains are represented. This is made up of various concepts that are technically and logically defined by relationship to more than one concept. To achieve consistency across the three domains, a shared terminology approach among all available sub-domains is essential.



Therefore, all three domains' information will be shared in the same architectural style which can be integrated in order to apply the uniform terminology to ensure interoperability among the healthcare, university and insurance domains. The proposed FDBS-based architecture resolves interoperability problem ranging from diversity of semantic and data structures adopted in the databases.

Each component of the federated database controls its interactions with other components by means of export and import schemas. The export schema specifies the information that a component will share with other components, while the import schema specifies the nonlocal information that a component wishes to manipulate (Frankfort-Nachmias, & Nachmias, 1992).

Each domain adopting a federated database system for achieving the interoperability (Vermeer & Apers, 1996) directs controls and monitors its relationship and interactions with other domains by means of import, export and private schemas as shown in Figure 16.

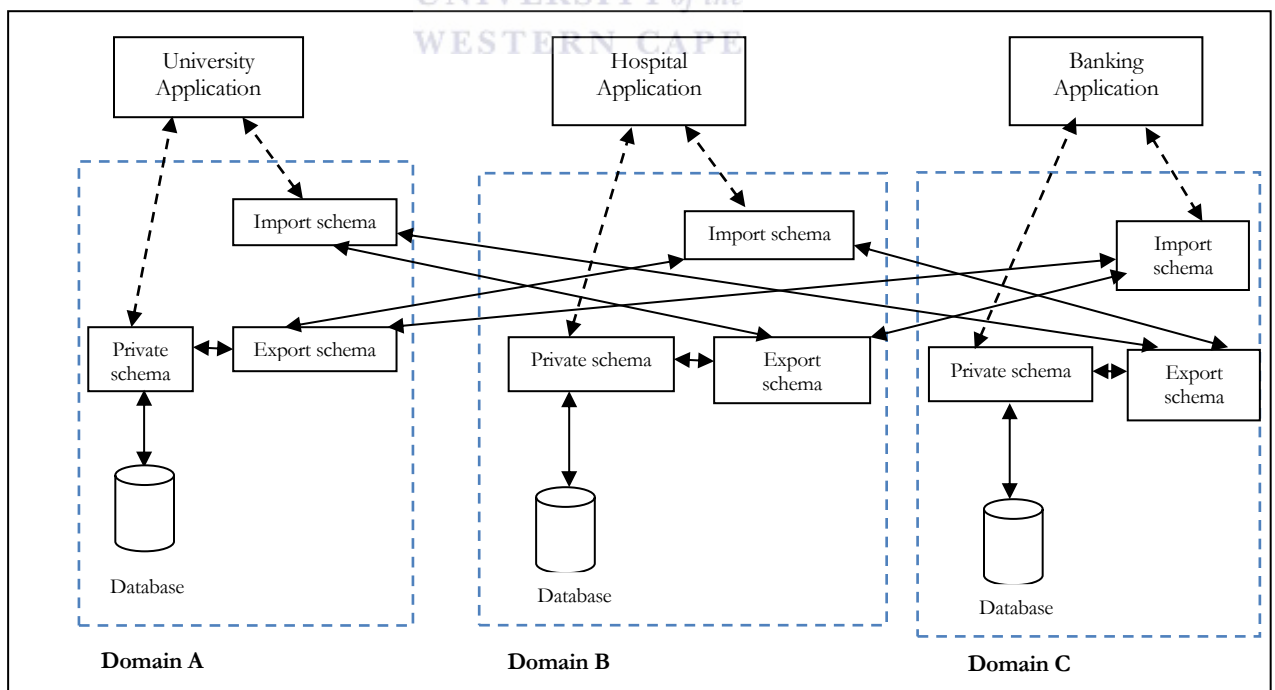


Figure 16: Federated database approach of achieving interoperability in a 3DBGE

*Declarative language (queries) for interoperability*

A major challenge to developing and implementing an interoperable 3DGBE lies in the ability to efficiently, sufficiently express cross-queries (inter-domain queries) that transfer information from different domains. To overcome this challenge, a declarative Meta-Query Language (MQL) (Vermeer & Apers, 1996) was adopted which is similar to the Structures Query Language (SQL).

MQL was used for querying and restructuring tables containing information across different domains. The main characteristic syntax of MQL is obtained directly from SchemaSQL (Lakshmanan, Sadri, & Subramanian, 2001), which is noted for higher-range declarations in the FROM clause. Another unique feature of MQL is its dynamic output relation which is specified in the SELECT clause. As illustrated in the scenario presented, MQL was used to query and restructure information across Domains A, B and C within the federated database. Hence, we achieved interoperability across these three domains through an MQL dynamic query mapping.

Some of the main characteristics of a language required for implementing interoperability as exhibited by MQL include the following (Laks, Fereidoon, & Subbu, 2001):

- i. The language must exhibit expressive power that is not dependent on the schema adopted by the database.
- ii. It must allow the rearrangement of a database so as to follow and conform to other schema.
- iii. It must permit efficient and effective implementation.

Three sample schemas are presented (See Appendix E) to describe interoperability using the template shown in Figure 16.

**CONCLUSION**

In this chapter, the research approach was presented vis-à-vis the epistemological, theoretical perspective, methodology and its related methods. Also, various methods adopted in addressing the challenges identified in this thesis were explored and

comprehensively discussed. Specifically, this chapter has presented the results obtained through content analysis, the approach used in designing the architectural framework, the method of simulation and the case study used for evaluation. Since part of this research work depends on simulation, a brief discussion on the choice of a simulator was discussed. Aside from the above, various definitions of some of the metrics used were defined. Finally, this chapter has successfully presented theoretical background as well as the approach used for implementing interoperability using a federated database approach. Summarily, this chapter has provided an insight into the results presented in chapter 4. In the next chapter, a detailed analysis of the results will be presented using the SSM as presented in this chapter.



## Chapter 4

### RESULTS

#### INTRODUCTION

In the previous chapter, the research approach was presented vis-à-vis, the epistemological, theoretical perspective, methodology and its related methods. In this chapter, the results of the evaluation and analysis are presented.

The chapter is structured according to the main sections of Figure 11. The first section covers content analysis, the second presents the system activity (architectural framework and algorithm design) and the third covers the simulation, experimentation and using the model to test its applicability with a case study.

#### SSM CYCLE 1: Content analysis

##### *A: Real-world situation of concern*

Scalability, interoperability and efficient access control have been identified as the major challenges confronting the full-scale adoption and utilization of grid computing. The effects of each of these have not limited its application alone, but have exposed the user to various security challenges resulting from inadequate access procedures in the VO.

##### *B: Purposeful activity: content analysis*

Grid computing challenges (scalability, interoperability and efficient access control) were identified as a concern and studies that address these aspects of grid computing were examined to determine commonalities among the reviewed literature. Content analysis was also used to establish a convincing argument regarding the significance of the research and where it leads, and to identify, any controversies relevant to the research. It was also used to bring to light any inconsistencies in findings relating to this area of study. Finally, it was used to identify any unanswered research questions.

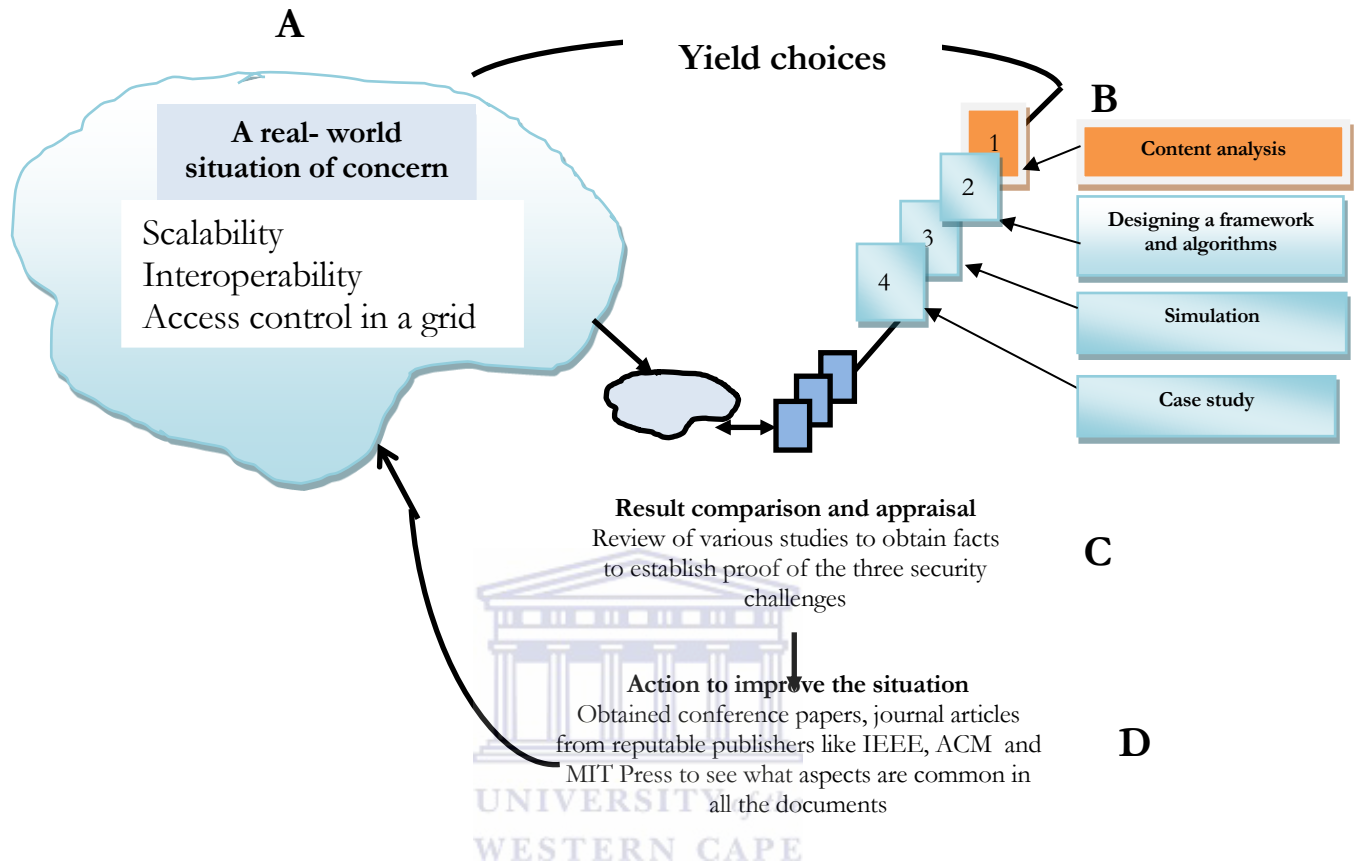


Figure 17: Content analysis method of identified grid challenges with SSM

Documentation regarding five security models were considered (see Appendix A), namely: access control list (ACL), mandatory access control (MAC); role based access control (RBAC), community based security model (CBSM), and discretionary access model (DAM). The consideration was based on a set of requirements: maturity, usage in the implementation of security policies, compatibility with various applications, and the security models' applicability on different platforms. The strengths and weaknesses of each were noted.

Furthermore, these models served as the basis for implementing and achieving access control, scalability and interoperability in distributed environments. Three aspects of these models were considered:

- implementation approach;
- technical strengths; and
- technical weaknesses.

All the models considered (ACL, MAC, RBAC, CBSM and DAM) have the same basic principles (see Figure 18). The user is an entity that demands resources through an “access request”, and the reference monitor provides a set of access requirements and access control policies to be fulfilled before any resource(s) can be accessed by an intended user.

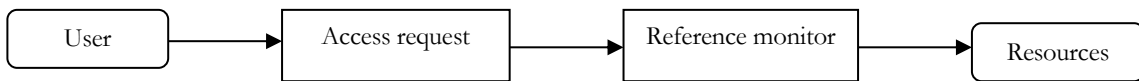


Figure 18: Basic principle of an access control model

Grid access control involves a passive object and an active subject, which has a particular access operation and a reference control (that either denies access or grants to resources or agents) (Higgins, Wilson, & Fell, 2005). On the grid, objects that are shared are various distributed resources (e.g. files, databases and supercomputers) and the subjects are the users and processes running on behalf of the various users. In most cases, access control is located at the user centre or resources centre (Baker, Buyya, & Laforenza, 2002).

Using content analysis, the implementation approach and strengths and weaknesses of each of the access control models were explored. The details of the context analysis are given in Appendix A.

### Summary of the five security models

Table 2 presents a summary of the analysis of the five security models as determined by the content analysis. (see Appendix A).

Table 2: Summarized features of five security models

Model	Applications	Strengths	Weaknesses
Access control list (ACL)	Constrains access to data and files on distributed systems	User privilege is clearly defined, a straightforward way of denying and granting access to resources is provided, and genuine users are specified for their resources	Lack of interoperability, no room for scalability and vulnerable to errors
Mandatory access control (MAC)	Allows every user and object to be assigned a label called a sensitivity label ,which comprises compartments and a level of secrecy	Handles intricate relationships among various entities in any computer-based environment and applications that involve sensitive data.	Difficult to implement, lack of interoperability, problem of blind write-up, information downgrading, not flexible and scalable
Distributed Authorization Model (DAM)	A user-centered authorization-oriented model translating a global credential to a local one	Full control by the local site director and administrator, allows for transparency, works with other models	Not flexible, not scalable, lacks interoperability
Context Based Security Model (CBSM)	Handles security challenges emanating from the high mobility of pervasive computing and different devices used in these types of environment	Enhances the reuse of policy specification, allows enforcement as well as fine-grained access control	New and requires further investigation, suffers from underutilization, and has poor access control mechanisms
Role Based Access Control (RBAC)	Administers individuals with roles, permissions and hierarchies in a dynamic and large-scale distributed environment such as a grid	Appropriate for administering various security policies; is flexible, scalable and interoperable; has a breadth of applications	It is being underutilised for handling current security and access control challenges in a distributed environment

*D: Action to improve the situation*

Since the primary objective of RBAC is to provide reliable access control of information in a VO and ensure security management across the grid network with the aid of role, services and permission specification for the user, RBAC was adopted for the development of the model framework. Figure 19 shows the statistical analysis of access control research and

papers reviewed from 2005 to 2012 and shows that the flexible RBAC model has been used for various applications due to its flexibility and ability to handle and solve the problems that emanated from the various traditional access control models.

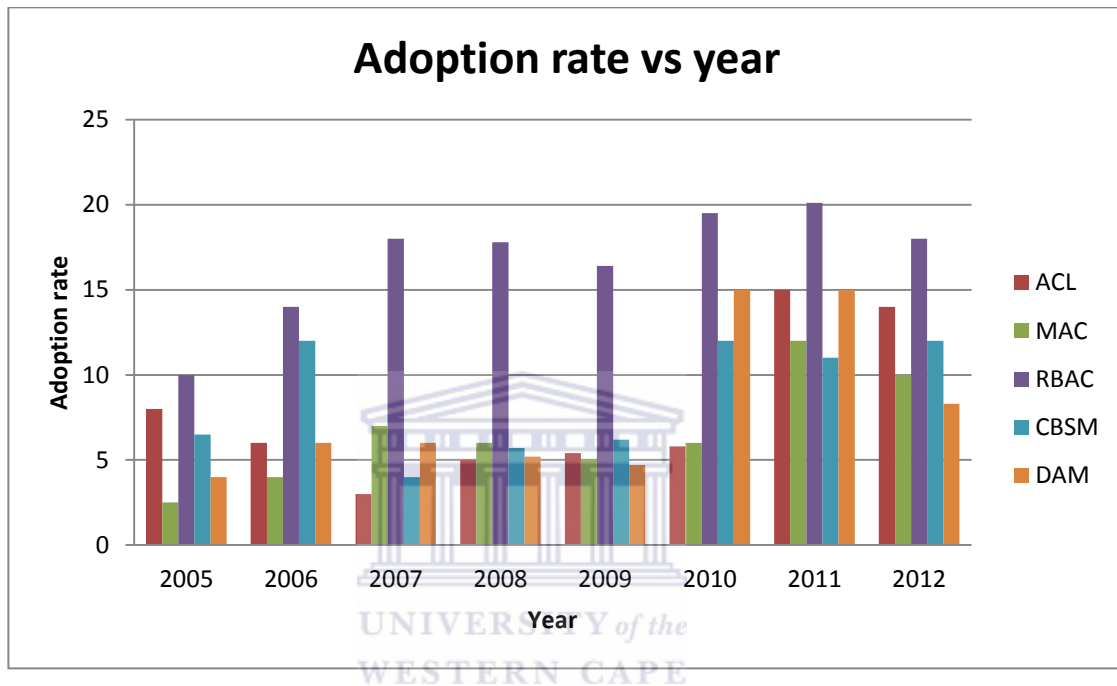


Figure 19: Adoption rate of the five security models. 2005–2012

Since 2006, RBAC had the highest rate of adoption of all security models. This was deduced after many applications, were observed in business, academe, administration systems, industry, the health sector, military operations and many more. Unlike other security models, RBAC has no boundary in its application since it is very flexible and allows room for both scalability and interoperability. Research carried out by Alessandro et. al. (2010) confirmed that RBAC is cost effective, minimizes complexity and reduces the cost of access permission management. These benefits have encouraged grid security architecture designers to adopt RBAC in managing access control on the grid and are the reasons for its adoption in this research project.



*Adoption rate* here is defined as the degree at which various researchers used a particular access control model for addressing various security and access control challenges.

### SSM CYCLE 2: Architectural design

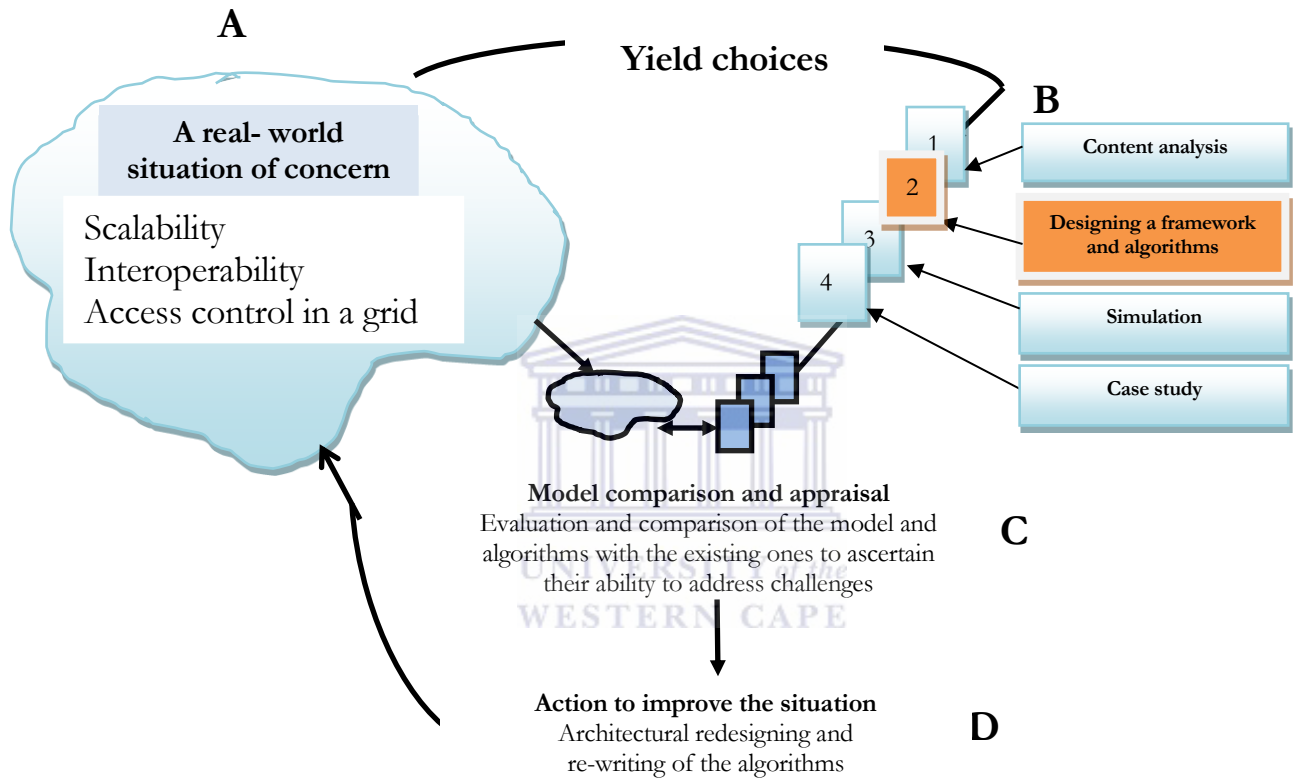


Figure 20: Designing the architectural framework and algorithm

*A: Real-world situation of concern*

RBAC was identified during content analysis as the most appropriate access control model for the grid. However, the security challenges, as previously explained, are still a concern. In this cycle, the results of the architecture and algorithm design are discussed in terms of the SSM (see Figure 20).

*B: Purposeful activity: architectural design*

*Strata of the proposed architecture*

The architecture is presented as two stages, each of which involves two phases (see Figure 21).





Phases involved in the proposed architecture			
	Grid environment	Fourth phase	} <b>Stage 2</b>
	Processing level	Third phase	
	CSMU	Second phase	} <b>Stage 1</b>
	Domains (LSMU)	First phase	

Figure 21: *Phases involved in the proposed architecture*

- i. The first phase involves various domains. Each of the domains is characterised by a user and an LSMU.
- ii. In the second phase, the CSMU interacts directly with all the domains of phase 1.
- iii. The third phase is a processing phase. All activities that result in the granting of resources are carried out in this phase.
- iv. The fourth phase is a grid environment phase where many resources are available. A user is allowed to access this phase based on a decision made in the third phase.

*Stage 1 of the architecture*

This stage involves the interaction between various users and the domains' LSMU with the CSMU. The architecture below gives comprehensive information with respect to this interaction and message passing between grid entities.

In Figure 22, a theoretical framework of the interaction between the user and the LSMUs of three domains, as well as its interaction of the three domains and the CSMU, is depicted. To

explain the process of the architecture presented in Figure 22, let us assume the following scenarios:

- i. Adam, a grid user in Domain A, forwards his request to his domain's LSMU, where his authorisation is verified and confirmed. Adam's status (eligibility as a user) is thus determined. This phase makes Adam's access right to the intended domain known.
- ii. The LSMU then sends Adam's request to access a resource in any intended domain to the CSMU to reconfirm his authorisation right in his own domain and his rights to access resources of any other domain. The CSMU verifies whether Adam qualifies to access the required resource. There are two outcomes: YES (acceptable) or NO (not acceptable).



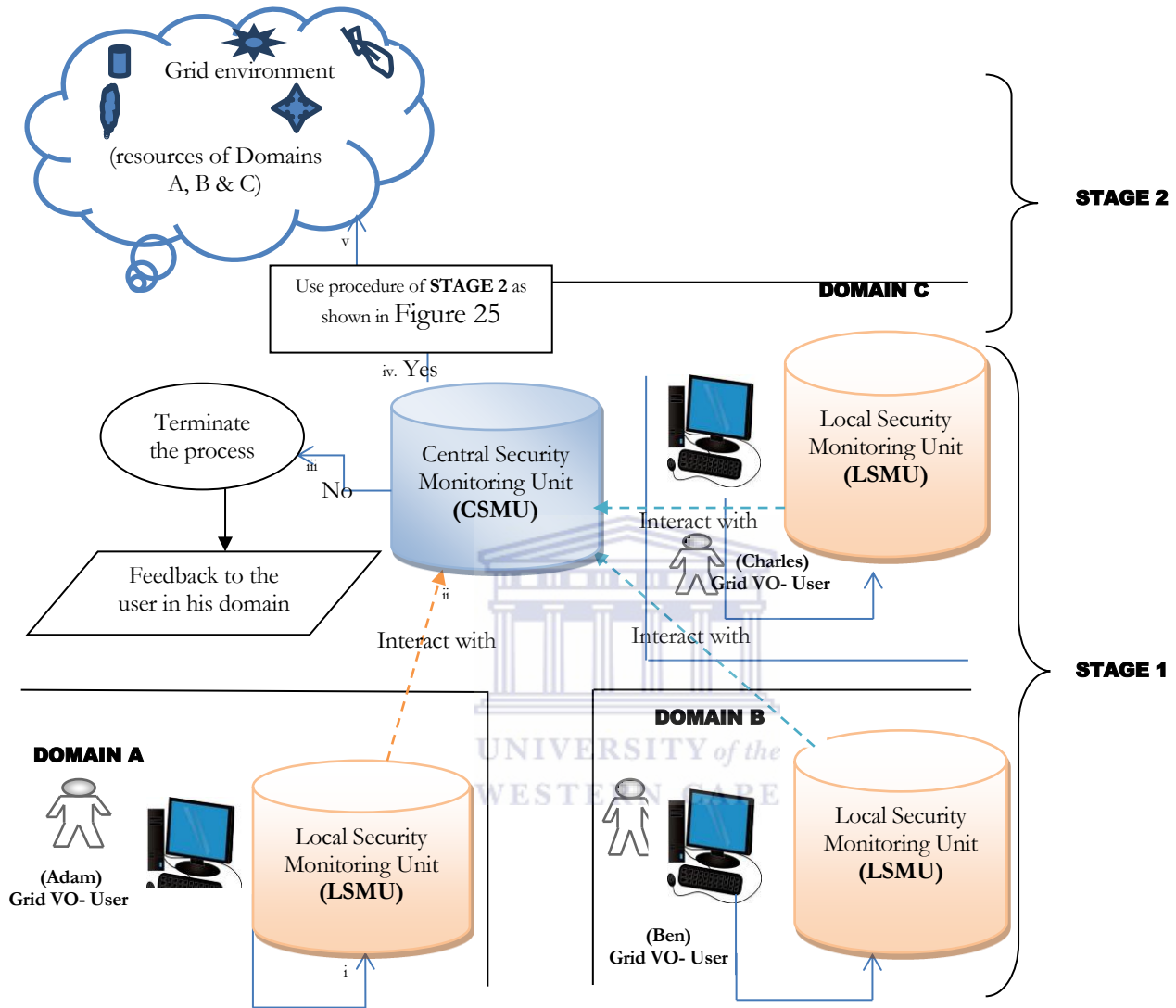


Figure 22: STAGE 1 of 3DGBE architectural framework of the proposed model (showing interaction between users, CSMU and LSMU)

- iii. If NO, the process (request) terminates and the feedback message is communicated to the user.

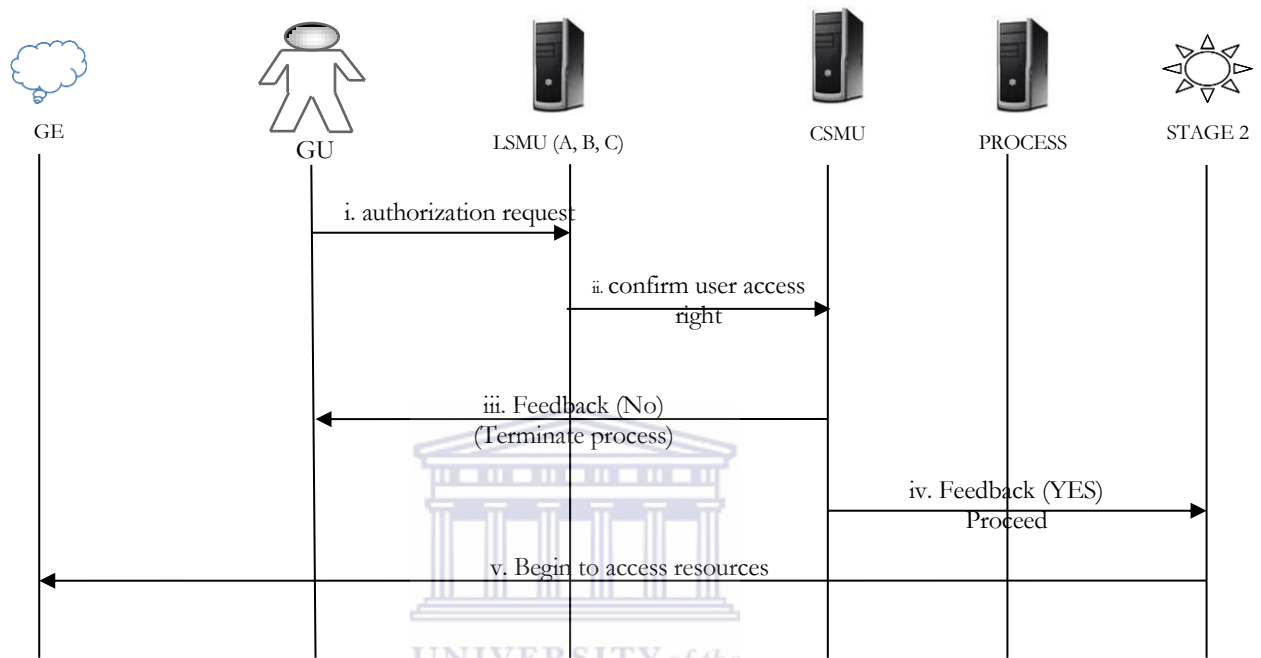


Figure 23: Information flow sequence of the architecture of Figure 22

- iv. If YES, a “clearance” certificate will be given to the user (Adam) by the LSMU of the intended domain and the user can proceed to stage 2.
- v. If there is a successful processing in stage 2, the user will proceed to access resources in the grid environment.

*Algorithm 1: Algorithm describing the working relation of components in Figure 22*


---

```

required : Domain A, Domain B, Domain C, LSMU, CSMU

begin:
feedback [authorisation] = "Yes or No";

GU{Domain A, B, C}: ← authorisationrequest LSMU:

                    If authorisation = "no"
                    then : terminate (process)

else:

                    if authorisation = "yes"

                        Then: LSMU → forwardrequest CSMU

CSMU → Verifies {(GU (role))}:

if CSMU [permission(decision)] = "yes"

    then: CSMU ← moves stage 2;

stop

```

---

Table 3: Functions of the architectural components

Component	Function
Policy access point (PAP)	Establishes and creates security policies and stores them in the proper security policy repository
Policy enforcement point (PEP)	Carries out access control by creating decision requests and also enforcing decision authorisation. PEP also assists in enabling quick and emergency policies as well as setting emergency attributes. It performs this by requesting access control decision and generating audit records.
Policy information point (PIP)	Stands as the only source and avenue for attribute values and/or the data that may be required for policy evaluation.
Grid user (GU)	The person on the grid that constantly makes a request from one security domain to another domain. His/her request in a VO is regulated by the LSMU (inside the domain) and CSMU (outside the security user's local domain).
The Grid User Authentication Service (GUAS)	Is responsible for providing a response to the request made for a privileged operation of a user before proceeding in accessing various resources on the grid
Policy decision point (PDP)	Examines the applicable and usable policies and renders the authorisation decision accordingly

*Stage 2 of the architecture*

This stage deals with the interaction between the processing phase and grid environment. This stage comes into play if and only if there is a positive feedback during stage 1 (see Figure 25).

Below is the operation of the architecture presented in Figure 25.

- i. Through the grid entry link, the GU requests access (with the role authorisation-certificate) from the GUAS. The request is either granted or not.
- ii. If the feedback is negative, the entire process will be terminated immediately and the request will cease to continue.
- iii. However, if the feedback is positive (YES), then the request will be forwarded to the PIP (a protocol of XACML for access control). This is to source detailed information about the user.
- iv. The request will further be directed to the PDP, which is another XACML protocol for access control. The PDP is responsible for making a decision on whether the user may access the requested domain. The feedback of the PDP will either be positive (YES) or negative (NO). If the feedback is negative, the entire process stops.
- v. If the feedback is YES the request is conveyed to the PEP.
- vi. The PEP will demand an updated version of the user permission certificate from the PDP (grid VO-PDP).
- vii. A certificate validation/update will be transferred to the centralized resource database server (CRDS) from the PDP (grid VO PDP).
- viii. Finally, a message will be sent to the user to proceed and access resources on the grid.

The procedure is applicable from either of the domains available on the grid i.e. either Domain A to Domain B or from Domain A to Domain C, and vice versa.

In order to ensure smooth and efficient access control on the grid and also to improve the performance of the architecture, the LSMU works with the CSMU. That is, there is smooth correspondence between the local security units of all the domains with the central security unit for the entire grid. They both communicate and work hand in hand to achieve a flexible, interoperable and scalable grid environment.

*Algorithm 2: Describing the working scenario of the architecture presented in Figure 25*

---

```

require: role, user, PIP, PEP, PDP, GUAS, CRDS
Feedback: [yes/no]:
  begin: from stage 1:
    request: —————>GU role certificate [GEL]
    then: GUAS <----- Role (GU);
           Verifies
           Else: if feedback (GUAS) = "No"
               Then: terminate (process);
           if feedback (GUAS) = "YES";
    Then: request —proceeds—> PIP;
    PIP: —request—>PDP; //request for appropriate decision
           If feedback = "yes"
           proceed: —————>PEP;
           Else if feedback = "No"
               then: stop (process)
    Getupdate:—————> PDP-VO;
    update (certificate): —obtained by—>CRDS;
    finalDecision: —Pass to—>(VO (grid))
    Begin [GU] :access [resource]
  stop

```

---

*Overview of the basic components of the architecture*

In the proposed model, each of the domains available in the VO has a LSMU that is responsible for the domain's local security access control and management. An advanced access control and management unit, called the CSMU handles access control and



authorisation interactions for the various grid entities across the three domains of the model. The CSMU within its capacity, along with the LSMU ensures interoperability, scalability, flexibility and secure access control for various grid entities across multiple administrative domains through inter-domain interaction; as well as application independence and its ability to accommodate additional grid entities.

For any access request by a grid user, the LSMU would verify the user's access privilege. The model is based on the adoption of the XACML's request-response protocol, which makes use of four basic components. The components are the PEP, PDP, PIP and PAP. However, in this model, only the PEP, PDP and PIP are used because of their relevance, usefulness and application in the proposed architecture.

#### *Basic assumptions*

The following were assumed:

1. A user from Domain A (Adam) may intend to access a resource in Domain B and a user in Domain B (Ben) may also be interested in accessing resources from Domain A.
2. A user in Domain A (Adam) may wish to access resources in Domain C, while a user in Domain C (Charles) may equally be interested in the resources of Domain A.

These are two possible scenarios when a three-domain-based architecture is being considered. Scenario 1 is explained in Figures 28 and 31 and is equally applicable to the other scenario.

Adam, Charles and Ben are users in the Domains A, B and C respectively. Each of them is bound by the security and access framework in their respective domains. There are six ways in which access could be requested:

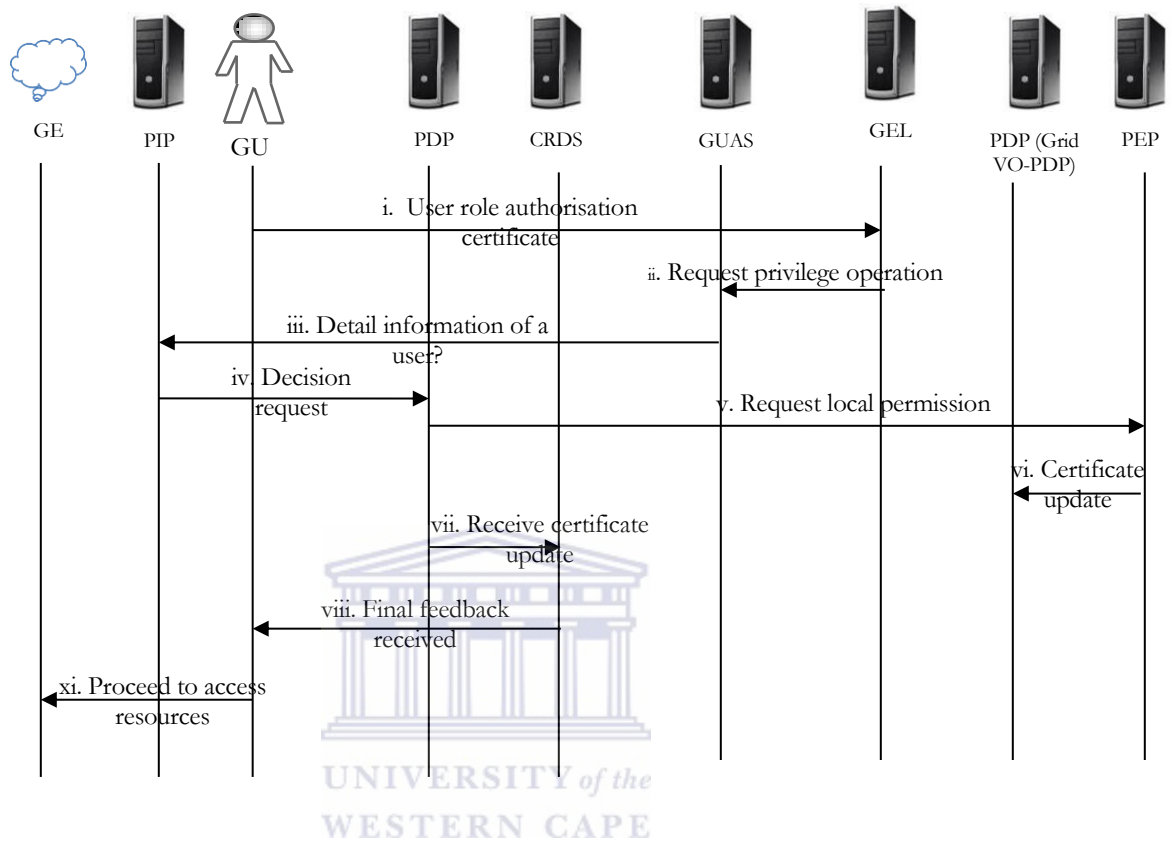


Figure 24: A 3DGBE stage 2 architectural framework information flow process of Figure 25

A → B; A → C; B → A; B → C; C → A or C → B.  
 Where “→” implies a request from one domain to another.

**Operational overview of the model**

The access control across three grid domains is complex and problematic. The security of the individual domains is quite dependable and efficient, because each of the domains has its own access control and monitoring policy, which is monitored by the LSMU. However If a user wishes to access resources in another domain, the user from the designated domain will first need to be verified by his domain. This is achieved by translating the certificate of his domain to the domain in which he wishes to access resources. The translation (or conversion) targets the access privileges and the identities in other domains on the grid. The CSMU is mainly in charge of monitoring and overseeing access and the security relationship

from one domain to another domain depending on where an entity is seeking access. Also, the CSMU is equally responsible for maintaining the information for the mapping of the interaction for the various domains. The process and the procedure for accessing resources across various domains are shown in Figures 28 and 31.

*C: Model comparison and appraisal*

Comparisons were made with respect to the architecture designed severally with the existing models. The newly developed architecture and algorithms were compared with the existing ones. The flaws and weaknesses discovered in previous designs are addressed in this research namely, lack of interoperability, inefficient access control and non-scalable grid based-environment. In summary, both the algorithms and the architectural framework were appraised based on the capacity to redeem and rectify the weaknesses of the existing models.

*D: Action to improve the model*

Both architecture and algorithms were improved upon by redesigning them in order to address the challenges that arose. The cyclical approach of SSM has assisted in continually improving the design in order to handle scalability, interoperability and access control challenges.

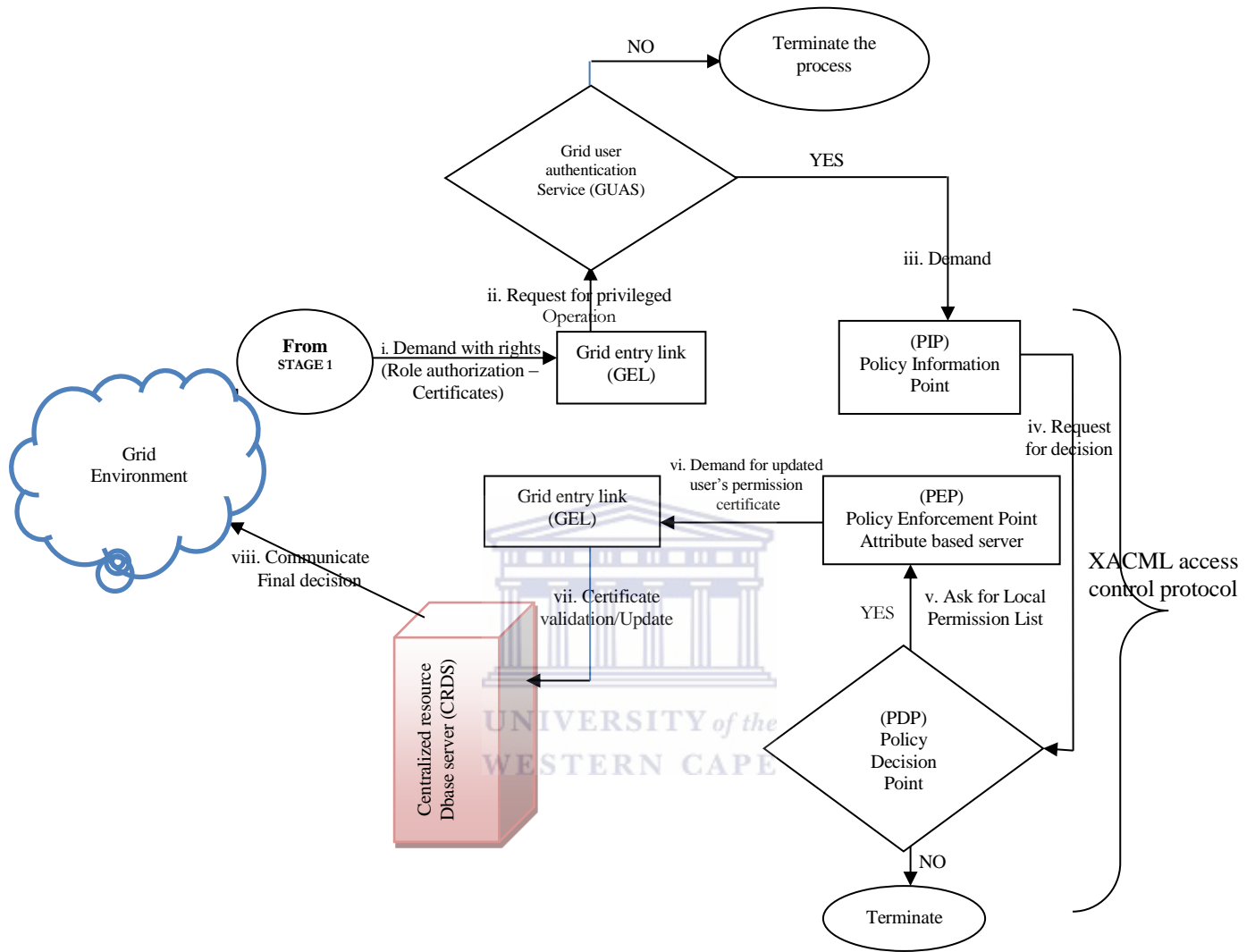


Figure 25: STAGE 2 of 3DGBE architectural framework of the proposed model

### SSM CYCLE 3: Simulation

#### A: Situation of concern

The simulation (see Figure 26) was carried out on the designed framework and algorithms in order to evaluate the performance of the architecture for addressing the challenges of scalability and interoperability.

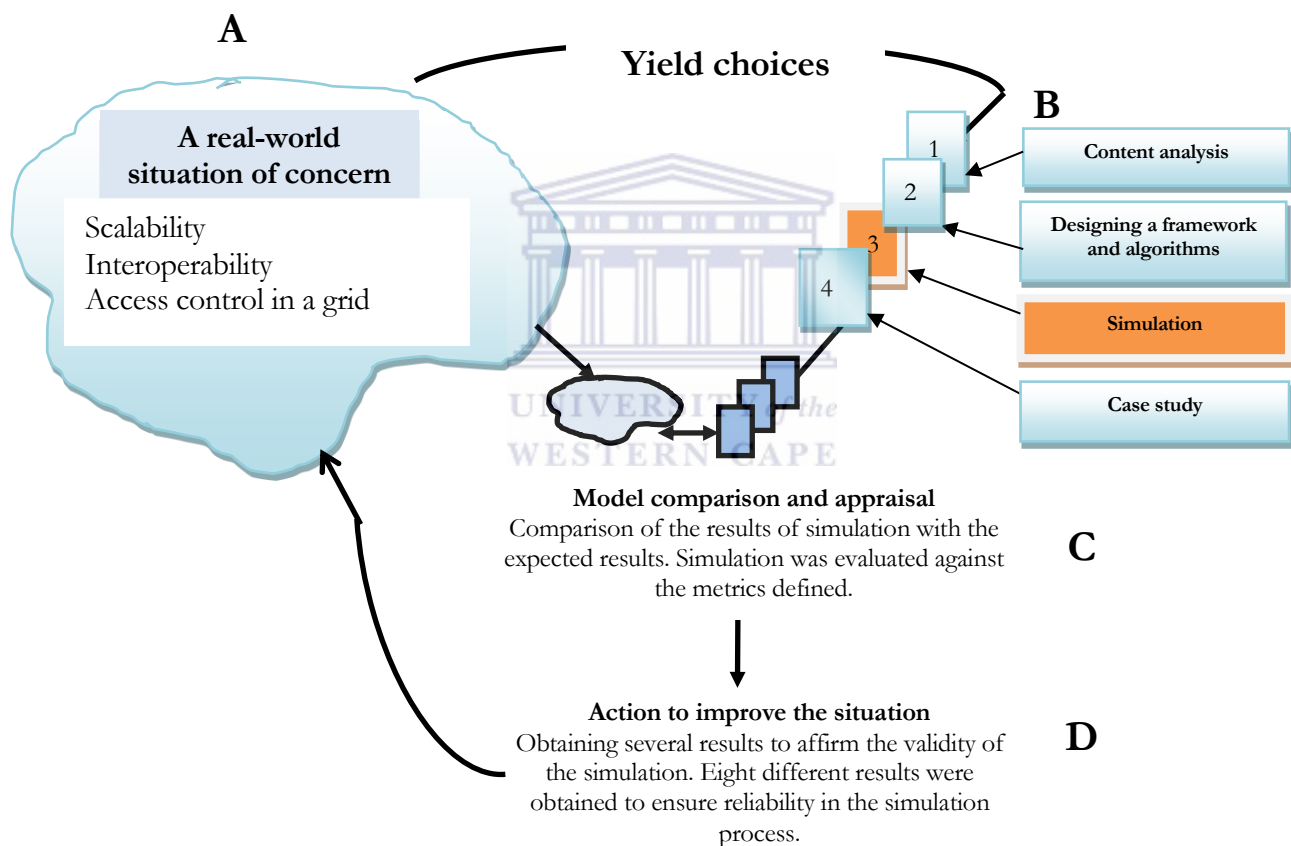


Figure 26: Simulation of designed framework and algorithms

#### B: Model comparison and appraisal

The results obtained with the metrics evaluated were compared with the expected results. The metrics being referred to are: average turnaround time and number of grid nodes;

average turnaround time and number of grid service requesters; throughput and number of grid nodes; number of access resources and the time. It was observed that as the number of grid nodes increased the average turnaround time reduced and thereby increased the number of service requesters (grid users) on the grid. (see Figure 29).

- The result also showed that as the number of service requesters increased, there was little effect on the turnaround time, which did not impact on the users' services and request time. (see Figure 30).
- It could also be deduced that as the number of grid node increased, the throughput also increased thereby increasing the number of resources being accessed within a given time. This proved the scalability of the 3DGBE architecture. (see Figure 31).

Several parameters were used in order to attain good simulation results.

#### *C: Action to improve the situation*

Eight different simulations were conducted for two metrics measured at a time for the experiments whose results are presented in Figures 29 and 30. The objective of this was to obtain valid, consistent and dependable results. The final results represented a valid outcome of the simulation. The mean, standard deviation and variance of the experiments whose results are presented in Figure 29 and 30 are calculated. (see Appendix F).

#### *First stage implementation: SCALABILITY*

Various simulation experiments have been carried out using different simulators. In this case, the Grid Security Services Simulator (G3S) was used. This is because; G3S is the only simulator that has security and access control functionalities. To carry out an empirical evaluation of the access control architecture, our simulation was developed in Java by using Jbuilder. There were three different domains in our experimental grid based environment: A, B and C. Domain A was made up of a cluster of computers that comprised seven nodes while the other two domains were LANs (local area networks) with 13 computers each. The

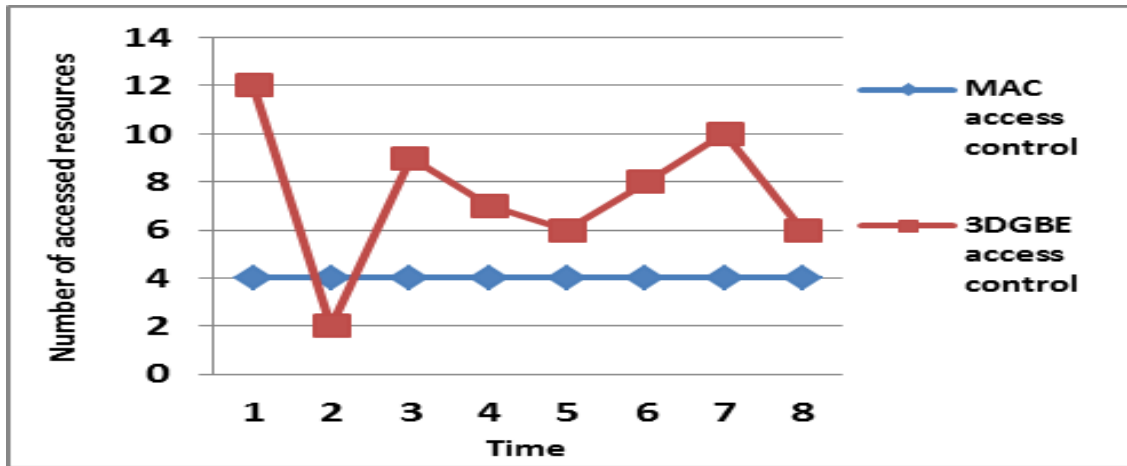


Figure 27: Number of available resources in two access control policies (3DGBE and MAC)

simulated grid environment was developed using the Globus toolkit 5.0.5. All the hardware of the testbed was embedded in Linux Ubuntu 12.04. A computer hosted a database with the information of all users and acted as the LSMU for each domain, while a computer server with a static IP address was chosen as the CSMU for the experimental grid. For efficient and reliable evaluation, we considered resources and entities that were accessible when a grid user requested their services. The result of our experiment revealed efficiency in terms of interoperability, scalability and efficient access control.

In the experiment, we compared 3DGBE access control and MAC, which is a popular access control method. Table 4 provides the detail of the parameters used in the simulation experiment. Users were provided with both a MAC-based and 3DGBE access control simultaneously. The number of resources was varied over different time periods. It was noted that the number of available resources varied over time in the 3DGBE access control architecture, whereas it remained unchanged in the MAC-based access control system; see Figure 27.

Comparing with MAC, 3DGBE access control performs tremendously better in terms of the number of accessed resources with a number ranging between 2 and 12 while the MAC remains at a constant of 4 which is below the average of 8 for 3DGBE access control. This means more flexibility of 3DGBE.

Table 4: Simulation parameters with their corresponding values

No.	Parameters	Corresponding values
1.	$\lambda_1$	0.25
2.	$\lambda_2$	0.36
3.	$DSR(a_i, \dots, a_i)$	0.34

Equation 2:  $DSR(A_i, A_j) = \frac{\sum_{a \in A_j} DSR(A_i, a)}{|A_j|}$  .....(2) was used to evaluate access without considering any weights. Entities in either Domain A, B or C would request resources from any destination and such requests would be evaluated by the destination domain. The result of SR was thereafter obtained. The result is shown in Figure 28.

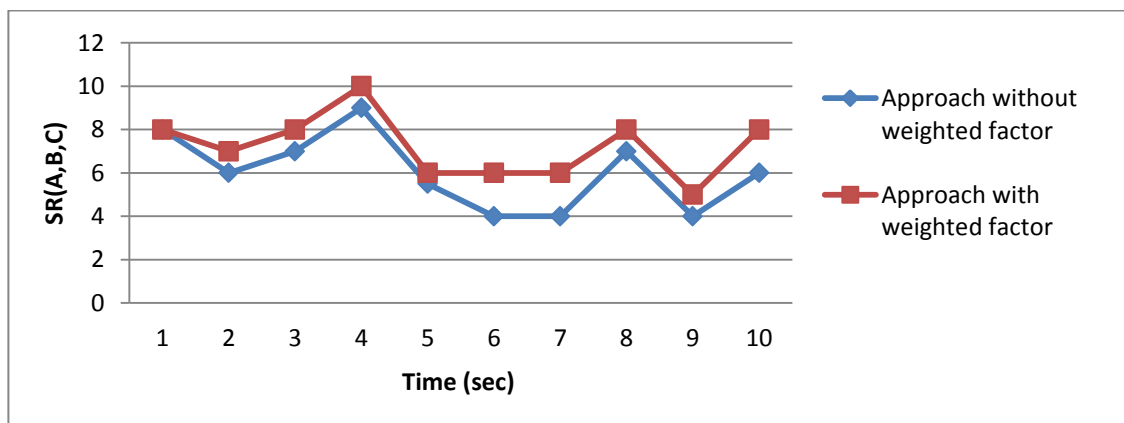


Figure 28: Secure rate comparison using two approaches



DSR and SR are two different defined parameters. DSR was defined for the evaluation of “Direct Security Rate” while SR was defined for “Security Rate”. Computable values were used and varied between the intervals 0-1.

Equation 1:  $SR(A, B, C) = \lambda_1 DSR(A, B, C) + \lambda_2 Rep(A, B, C)$  .....(1) was used for calculating the SR between the domains. The security rate value will vary if there are no weighted values for  $\theta_j$ . The secure rate of the approach with weighted factor seems to be better as projected than that without weighted factor. Table 5 gives a summary of the required parameters.

In order to arrive at the final result as presented in Figure 28, the values of parameters used ( $\lambda_1$ , value of DSR(A,B,C), value of Rep(A,B,C), entities in Domain A, entities in Domain B and entities in Domain C) were consistently varied. Different results were also obtained. This was done so as to reliably yield the same and comparable result with varied values.

Table 5: *Simulation parameters for  $\lambda_1$ , DSR, Rep for Domains A, B, and C and the number of entities of each domain*

No.	Parameters	Corresponding values
1	$\lambda_1$	0.6
2	First (initial) value of DSR(A,B,C)	0.58
3	First (initial) value of Rep(A,B,C)	0.44
4	Entities in Domain A	20
5	Entities in Domain B	15
6	Entities in Domain C	23

The direct relation between the number of grid nodes available and the turnaround time has been established and the nodes have direct influence on the time.

Three different machines were set up for Domains A, B and C enabled with Globus 5.0, gLite and Alchemi middleware as well as Unix, Linux and Windows operating systems

respectively. The turnaround time obtained for each of the domains based on the resources required when the number of grid nodes was varied. The average turnaround time for the entire grid was plotted against the number of grid nodes. The results are presented in Figure 29 and 30. The mean, standard deviation and variance were calculated using ANOVA (see Appendix F).

The simulation result revealed that the number of grid nodes that are available has a direct influence on the turnaround time, as shown in Figure 29.

This implies that as the number of grid nodes increases the average turnaround time reduces, and thereby increases the number of service requesters (grid users) on the grid. To further prove the argument that the model developed and implemented is scalable, Figure 30 shows that as the number of service requesters increases, there is a slight effect on the turnaround time, which does not impact on the users' services and request time.

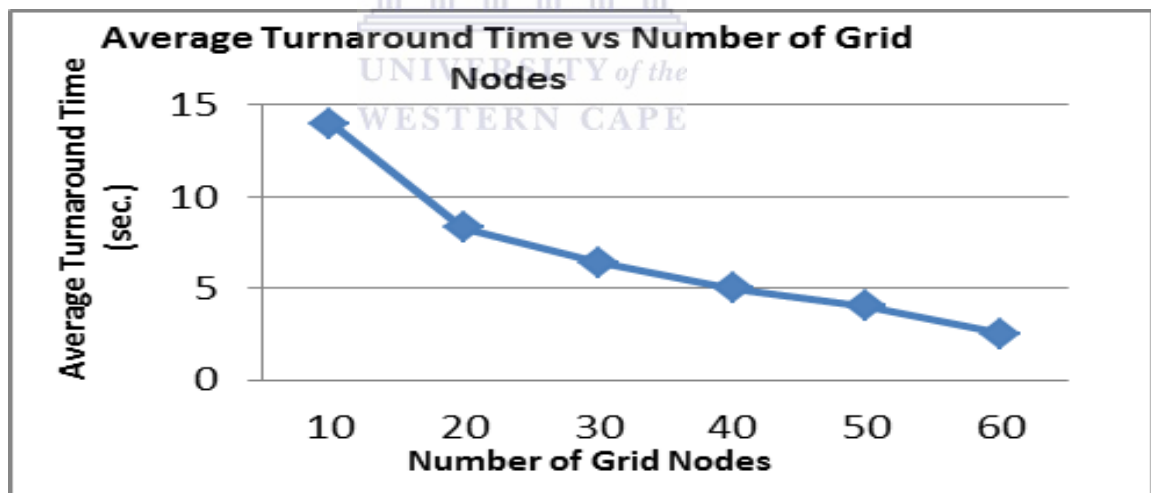


Figure 29: Average turnaround time versus number of grid nodes

A scalable architecture should give room for expansion and increase in network and resources through additional grid nodes and users. This is to allow additional resources and

users on the grid network. For any architecture to be considered as scalable, it should be able to meet both present and future throughput.

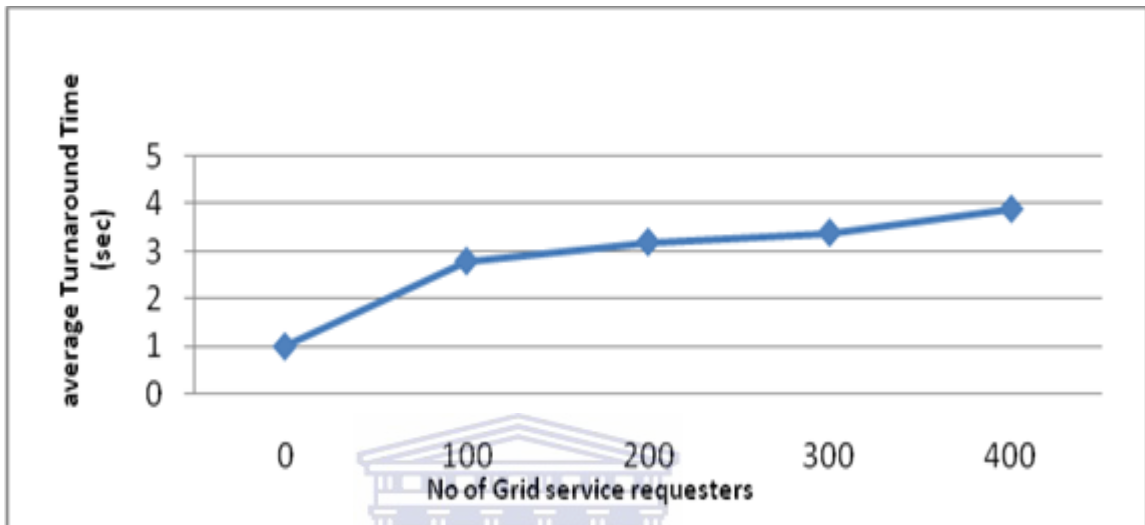


Figure 30: Average turnaround time versus number of service requesters

In order to prove that the 3DGBE architecture is scalable, we observed and measured the effect of an increase in the number of nodes against the volume of data that is transferred within a given period of time (throughput).

The initial setup indicated that Domains A, B and C contain 7, 13 and 13, nodes respectively. To ascertain the effect of increase in nodes on the performance of the *throughput*, the number of nodes in each domain was increased as follows: Domain A to 12 nodes; Domain B to 20 nodes and Domain C to 25 nodes.

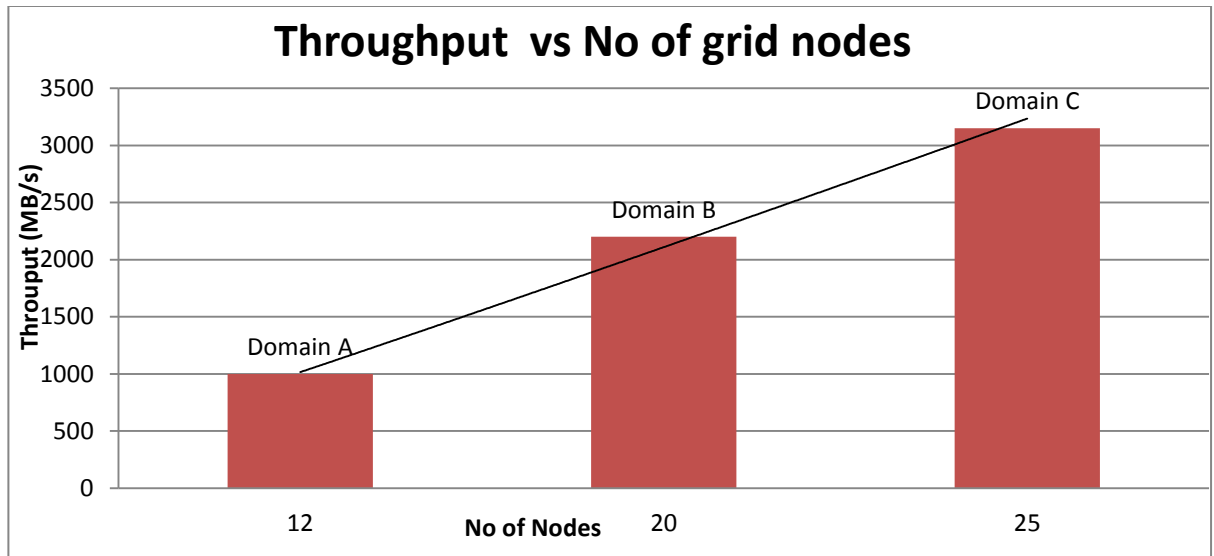


Figure 31: Throughput (MB/s) vs number of nodes

To study the performance of nodes against throughput, two experiments were conducted on two different machines already deployed with G3S. For the first machine, a Windows operating system enabled with GLite middleware was used. From the hardware's point of view, Pentium-4 (P4), 4.5 GHz processor was used with 4 GB RAM.

The second experiment was carried out on a machine where Linux operating system was installed with Globus 5.0. The hardware is made of Intel i3 processor + mother board 7000rs, 5GB RAM 850rs and 500 GB H-disk 1700.

Having achieved the results concerning the measurement of effect of grid nodes on the throughput, it is worthy of mentioning that these seemingly interesting results were also affected by the configuration of each of the machines. Most importantly, Globus 5.0 runs faster on Linux operating system and thereby produces greater values on the return of the turnaround time as depicted in Figure 32. Conversely, GLite runs slowly especially when enabled with windows operating system.

From Figure 31 it can be deduced that as the number of *grid node* increases, the *throughput* also increases, thereby increasing the number of resources being accessed at a given time. This further proves the scalability of the 3DGBE architecture. The result obtained as shown

in Figure 31 shows an increase in throughput as follows: when the number of grid nodes in Domain A is 12, the throughput is 100MB/s, when the number of grid nodes in Domain B is increased to 20, the throughput is 2200MB/s, while 3100MB/s is attained when the number of grid nodes in Domain C is increased to 25.

Conversely, the effect of decreasing in grid node across domains was measured in terms of throughput and the result is presented in Figure 32.

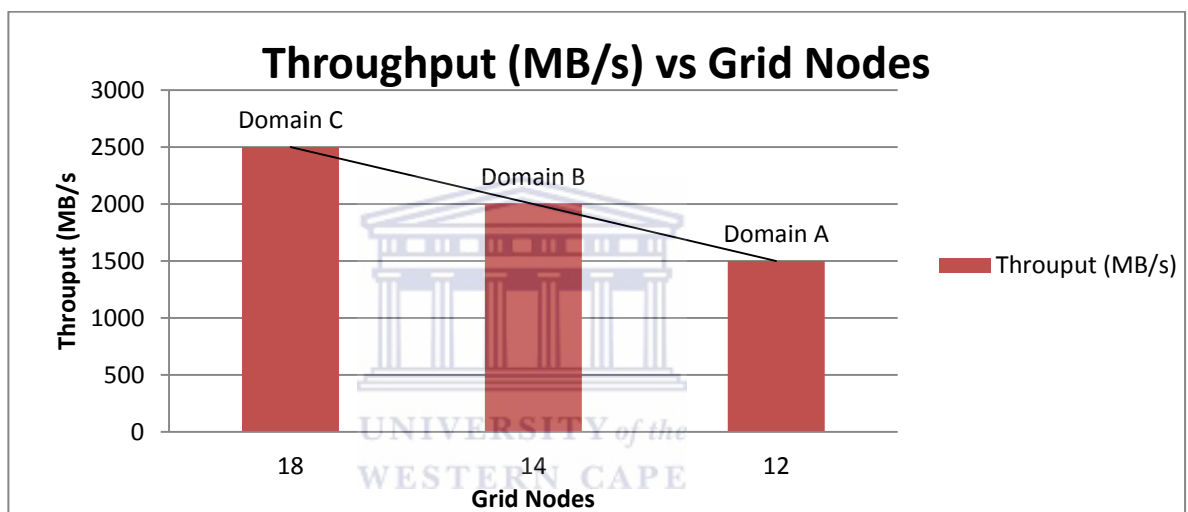


Figure 32: Decreased effect of grid nodes on throughput

The result obtained indicates the gross decrease in the throughput due to the reduction in the number of grid node across the three domains on the grid.

The results obtained in Figure and 32 therefore seem inconsistent in terms of the expected throughput due to the hardware and software configuration of the two machines.

#### *Second stage implementation: INTEROPERABILITY*

Interoperability can be defined as the capability of various domains to bring together resources and operate in a collaborative manner without any constraint. Interoperability plays an important role in resource sharing in a multi-domain environment. A grid-based environment that is not interoperable poses a security challenge to resources across multiple

administrative domains as users scramble to access resources from domains other than theirs without success. An effective interoperable VO requires sharing interactions among potential grid participants and, as such, interoperability in terms of people, services and various resources is one of the main issues addressed.

To this end, in order to achieve efficient interoperability in a 3DGBE environment, four basic approaches were focused on, namely:

- a. Grid middleware services;
- b. Appropriate handling of authentication and authorisation through LSMU and CSMU;
- c. Operating system interoperability; and
- d. A federated database.

*Interoperability with grid middleware (using tri-middleware integration in 3DGBE)*

Aneka, Alchemi (Luther, Buyya, Ranjan, & Venugopal, 2005), Cosm P2P Toolkit, Globus, Gridbus, Grid Datafarm, GridSim (Toolkit for Grid Resource Modeling and Scheduling Simulation), Jxta Peer to Peer Network, Legion, NorduGrid middleware (Buyya & Venugopal, 2004), PUNCH, Simgrid, Storage Resource Broker (SRB), ProActive, Unicore, and Vishwa are prominent among the grid middleware (Buyya, 2002). With some of the listed grid middlewares, the VO interoperability issue remains unabated. This is because of the absence of upper-level semantic concepts on their (grid middleware) layers (Welch & Lathrop, 2003).

To address this challenge, a tri-middleware integration approach was used (see Figure 33)

Middleware can be regarded as:

*“A mediator layer that provides a consistent and homogeneous access to resources managed locally with different syntax and access methods”* ( Priol, 2008, p. 32)

As shown in Figure 33, the 3DGBE was enabled with three different middlewares across the three available domains on the grid. The middlewares used were Globus 5.0, gLite and Alchemi for domains A, B and C, respectively.

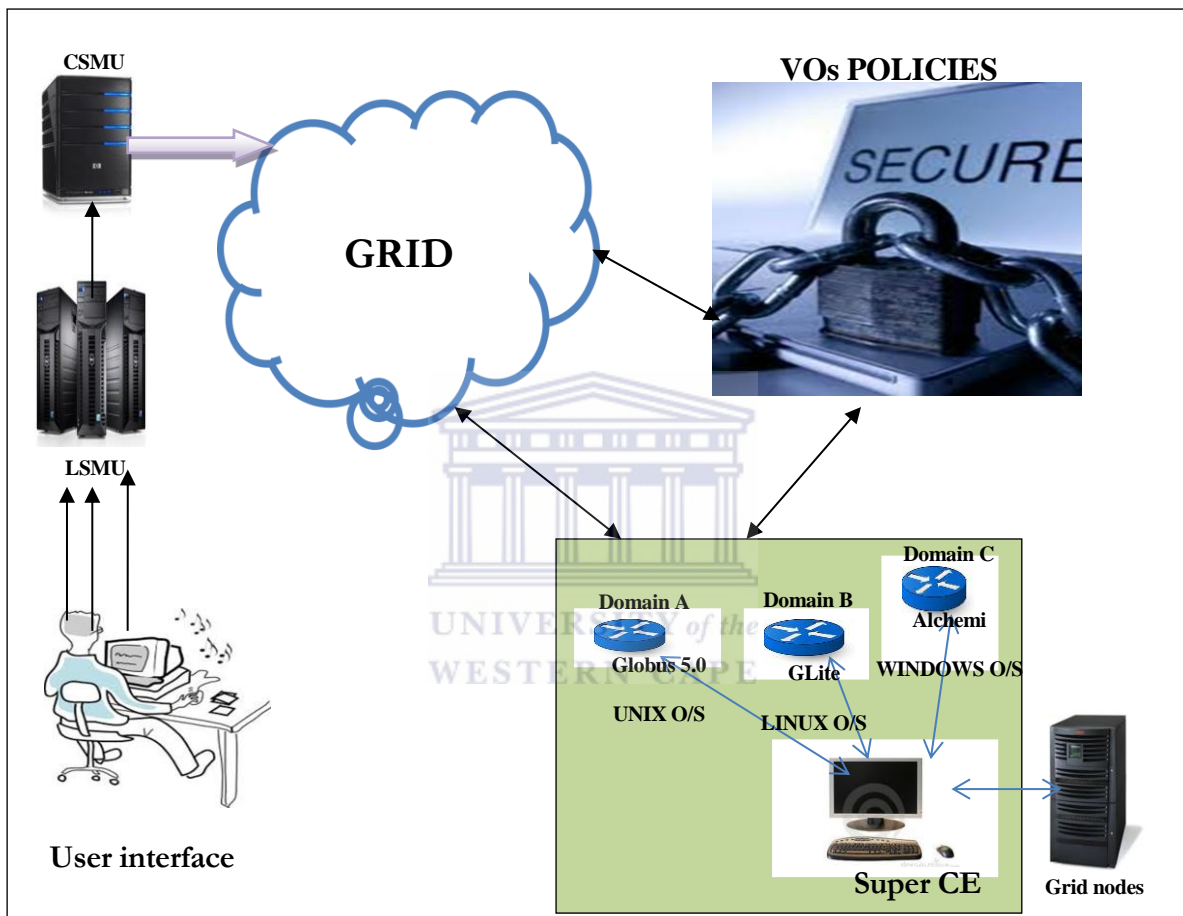


Figure 33: Tri-middleware integration based infrastructure for 3DGBE interoperability

With this approach to grid interoperability, which is based only on middleware integration, various middlewares were deployed on different domains to allow the same set of users to share and access resources with well-established and defined virtual organisation policies, irrespective of the grid middleware they intend to use.

The problem of middleware differences was solved by the effective usage of common standards that also take into consideration all the features of each middleware, which were arranged as a subset of specifications that needed to be implemented in different grid middleware.

In this approach to interoperation based on middleware integration, different middleware do not necessarily need to communicate with each other in order to have three or more grid infrastructures based on them to be able to merge and share resources.

Various computing resources currently installed in the three domains were all made accessible to all the grid users independently of the middleware they intended to adopt. The concept is depicted in Figure 33.

Consequent upon the implementation of middleware integration across the three domains, the aggregation of three or more grid resources was easier and the grid user in any of the three domains could decide where and when to access resources without hindrance.

Aside from Globus 5.0.2, the other two middleware installed for the interoperability integration testing were gLite and Alchemi. The choice of these three middleware was due to their availability (diffusion) and for technical reasons. Their availability has to do with the presence and relevance of these middleware in South Africa while the technical factor has to do with the variation in their architectural models and the likelihood of integrating every component into a uniform entity.

Basically, local clusters and security are considered as the basic elements of infrastructure that could be reused to aid interoperability across the three domains. Globus 5.0.2, gLite and Alchemi aid Torque/MAUI as the local scheduler that supports the ability to share local clusters and all available resources. Also, all these middleware use X.509 certificates, hence the same grid resources can be accessed, shared and distributed by these three different grid middleware.



Achieving interoperability with three different middleware is simple with Globus 5.0.2, gLite and Alchemi. Both gLite and Alchemi adopted the grid system infrastructure (GSI) model developed by Globus for user authorisation. This model (GSI) makes use of digital certificates and proxies for the authentication and authorisation of hosts and users. Established on X.509 digital certificates and proxies, GSI was extended in both the gLite and Alchemi (Rahman & Davis, 2009) with the agreement of the Virtual Organisation Membership Service (VOMS), which releases fully X.509 compatible signed extensions to proxies. Additional information about users that is required for the mapping on various levels of authorisation is achieved through these extensions. Since VOMS proxy is compatible with X.509 proxy, the former's proxy can be taken as authentication and authorisation credential when deploying on the three grid middleware.

The distribution of resources across a tri-middleware based architecture is the second method of achieving interoperability. The cluster manager in charge of the local resources was configured in such a way that jobs can be submitted despite differences in middleware. The local scheduler in architecture is Torque/MAUI. This scheduler is supported by Globus 5.0.2, gLite and Alchemi, hence it is very easy to express new queues new and added middleware in order to utilise the same resources.

#### *Appropriate handling of authentication and authorisation through the LSMU and CSMU*

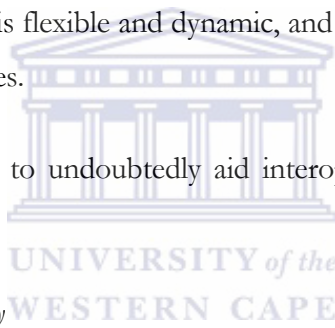
The CSMU maintains a high degree of interoperability between the users on the grid. For any resource request to be allowed, after approval has been given by the local security monitoring unit (LSMU), the former serves as the central nerve through which the final decision is made to permit grid resources to be accessed. There is smooth correspondence between the LSMU of each of the three domains and the CSMU. The purpose of this is to ensure an adequate and efficient data-sharing mechanism among the domains with a view to achieving interoperation through authorisation. The CSMU forwards the request of a grid user confirming his/her authorisation to access resources.

This new authorization and authentication approach also addresses the challenges of existing grid systems' access control procedures. Firstly, using a centralized authorization server (a CSMU) to manage the access control policies within an organization unit range greatly minimizes the administrative burden on each individual domain.

Secondly, at the application level, a grid user can check his/her authentication and authorization information details prior to sending access requests to various domains. This makes resource allocation processes within the grid system more effective, reliable and efficient.

Thirdly, by applying RBAC to specify access control policies on the authorization server, the authorization approach is flexible and dynamic, and provides fine-grained access control privileges to system resources.

All these features combine to undoubtedly aid interoperability across the three domains existing on the grid.



#### *Operating system interoperability*

To investigate interoperability for 3DGBE, Tivoli Storage Manager (TSM), was used across the three domains on the grid. Domains A, B, C were embedded with UNIX, LINUX and Windows operating systems, respectively (see Figure 33) Interoperability was achievable across the domains by setting up and establishing a consistent naming convention and strategy across the domains. TSM is a dedicated and centralised enterprise class recovery package and policy-based software that allows various users to share, collaborate and integrate across domains irrespective of the operating system used.

Aside from the fact that the architecture permits various applications to run (application interoperability), the architecture has proved to be interoperable on the UNIX, LINUX, (Ubuntu 10.04) and Windows operating systems that were deployed with different middleware (Globus, Glite and UNICORE).

A comparative evaluation of 3DGBE was carried out against the existing access control models (MAC, CAS, AKENTI and PERMIS) by providing a uniform number of resources for access within a given period of time on a TSM API. Efficiency was measured by calculating the ratio of accessible resources within a given time frame.

In comparing the level of interoperability of 3DGBE with other access control models, the efficiency was determined by measuring the ratio of resources accessed with the response time. This was achieved when UNIX, LINUX (Ubuntu 10.04) and Windows operating systems were deployed with Globus, Glite and UNICORE middleware. Table 6 gives values obtained.

Table 6: Comparative evaluation of interoperability of 3DGBE with other models

User	Models	No of resources accessed	Response time (sec)	Efficiency ( $\eta$ )
A	MAC	60	120	0.5
B	3DGBE	130	25	5.2
C	CAS	101	89	1.13
D	AKENTI	98	45	2.2
E	PERMIS	34	56	0.61

The efficiency denoted as ( $\eta$ ) is a measure of the ratio of the number (no.) of resources accessed (retrieved) to the response time (sec) using a specific model.

1. Efficiency ( $\eta$ ) for **MAC** = No. of resources accessed / Response time (sec)

$$= (60 / 120)$$

$$= 0.5$$

2. Efficiency ( $\eta$ ) for **3DGBE** = No. of resources accessed / Response time (sec)

$$= (130 / 25)$$

$$= 5.2$$

3. Efficiency ( $\eta$ ) for **CAS** = No. of resources accessed / Response time (sec)

$$= (101 / 89)$$

$$= 1.13$$

4. Efficiency ( $\eta$ ) for **AKENTI** = No. of resources accessed / Response time (sec)

$$= (98 / 45)$$

$$= 2.2$$

5. Efficiency ( $\eta$ ) for **PERMIS** = No. of resources accessed / Response time (sec)

$$= (34 / 56)$$

$$= 0.61$$

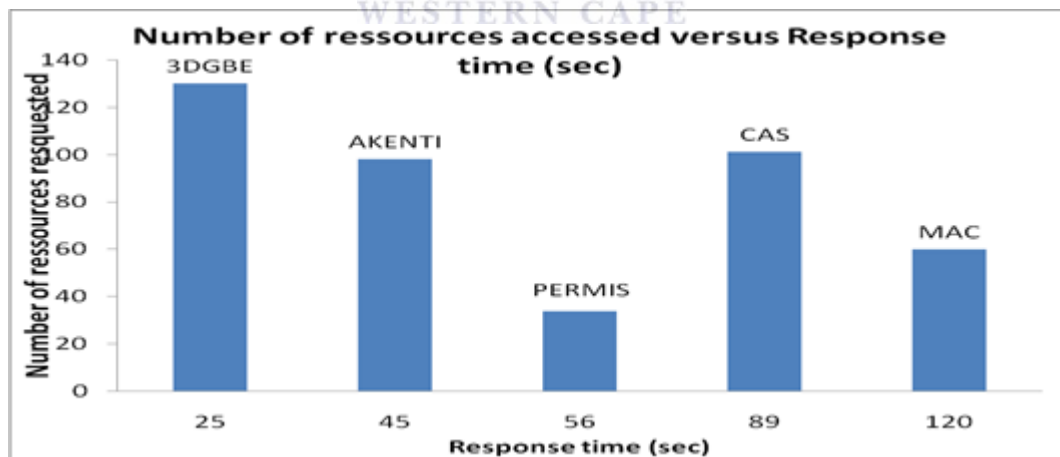


Figure 34: Comparative evaluation of interoperability of 3DGBE with the existing system

To buttress the fact that 3DBGE is interoperable, a comparative analysis was made with other prototypes currently available and the result is provided in Figure 34.

#### *A federated database: data interoperability in a 3DGBE*

##### *Implementation*

For the sake of explicitness and clarity, to implement interoperability in 3DGBE using the approach depicted in Figure 15, we created different databases for each of the domains. A university database, hospital database and banking database were created for Domains A, B and C, respectively. They are depicted in Tables 10, 11, 12, 13, 14, 15, 16, 17 and 18 respectively.

The information of all three domains was interrelated. Ensuring interoperability therefore did not only allow information sharing across domains, but prevented data redundancy across the three levels of architecture.

Efforts made at ensuring that the three domains become interoperable for information involved data integration, which involved aggregating and combining sources of data in all domains, i.e. multiple, distributed sources and locations to give a single and central database using a federated database system. This involved integrating different applications across varied platforms such as Windows and Linux operating systems. It also involved developing libraries of schema to ensure data sharing.

##### *Test scenario*

Three different databases were created for Domains A, B and C, namely a university, hospital and a banking database respectively. The university database comprised four tables (see Tables 10-14 in Appendix B), while the hospital has three tables (see Table 14-16 in Appendix B) and the banking database two tables (see Tables 17-18 in Appendix B).

##### *Inter-domain queries*

In order to achieve interoperability from the database created for the domains, the following approaches were adopted:

- i. A query was issued for databases in Domains A and B and the corresponding result was obtained.
- ii. Another query was issued for databases in Domains A and C using Metal-Query Language.
- iii. Finally, a query was issued for Domains A, B and C and the result obtained were presented accordingly.

**CASE 1:** two different databases were created for Domains A and B and they needed to be aggregated and joined to enhance further operations through queries. To achieve this, the query below was issued and the result obtained showed that the databases in Domains A and B were joined together for further operation. The query is given in Figure 41.

```
SELECT DB1.*, DB2.*
FROM DomainA.Database1.dbo.myTable AS DB1
INNER JOIN DomainB.Database2.dbo.myTable AS DB2
ON DB1.id = DB2.id
```

Figure 35: Query for aggregating data from Domains A and B

**CASE 2:** in an attempt to combine all the information on the databases of Domains A and C, the following query was issued and the report generated shows a UNION of both databases for both domains. The query is given in Figure 36.

```
-- FROM Domain_A
SELECT *
FROM [MyDatabaseOnDomain_A].[dbo].[MyTable] Table1
INNER JOIN
[Domain_C].[MyDatabaseOnDomain_C].[dbo].[MyOtherTable] Table2
ON Table1.ID = Table2.ID
```

Figure 36: Query for aggregating data from Domains A and C

```
1. SELECT a.Student_NO, a.Surname,a.Nationality,  
2. b.File_No,b.Patient_Condition, b. Patient_Account,  
3. c.Service_Code, c. Account_ID, c.Tax_ID,b.Age  
4. FROM Table16 c,Table13 b,Table9 a  
5. ;
```

Figure 37: Cross-domain query for joining data from Domains A, B and C

The cross-domain queries were then applied to each of the newly obtained tables. To obtain IDNO, Surname, and Nationality from Table 8; File No, Patient\_Condition and Age from Table 8; as well as Service\_Code and Tax\_ID from Table which belong Domains A, B and C respectively, the following condition holds:

- “a” , “b” and “c” are used as the aliases for Table 11, Table 15 and Table 18 respectively. Hence the following cross-domain query was used. The result obtained (see Appendix C) shows the possibility of integrating data from various domains.

Aside from the above queries, some other queries were issued to "SELECT" and "JOIN" databases across domains. The cross-domain queries were introduced purposely to handle the heterogeneity of information represented in different structures and to provide distinct aggregation capability in addition to the principal objective of multi-domain database interoperability. (See Appendix C for the report generated on the queries.)

*Third stage of implementation: EFFICIENT ACCESS CONTROL*

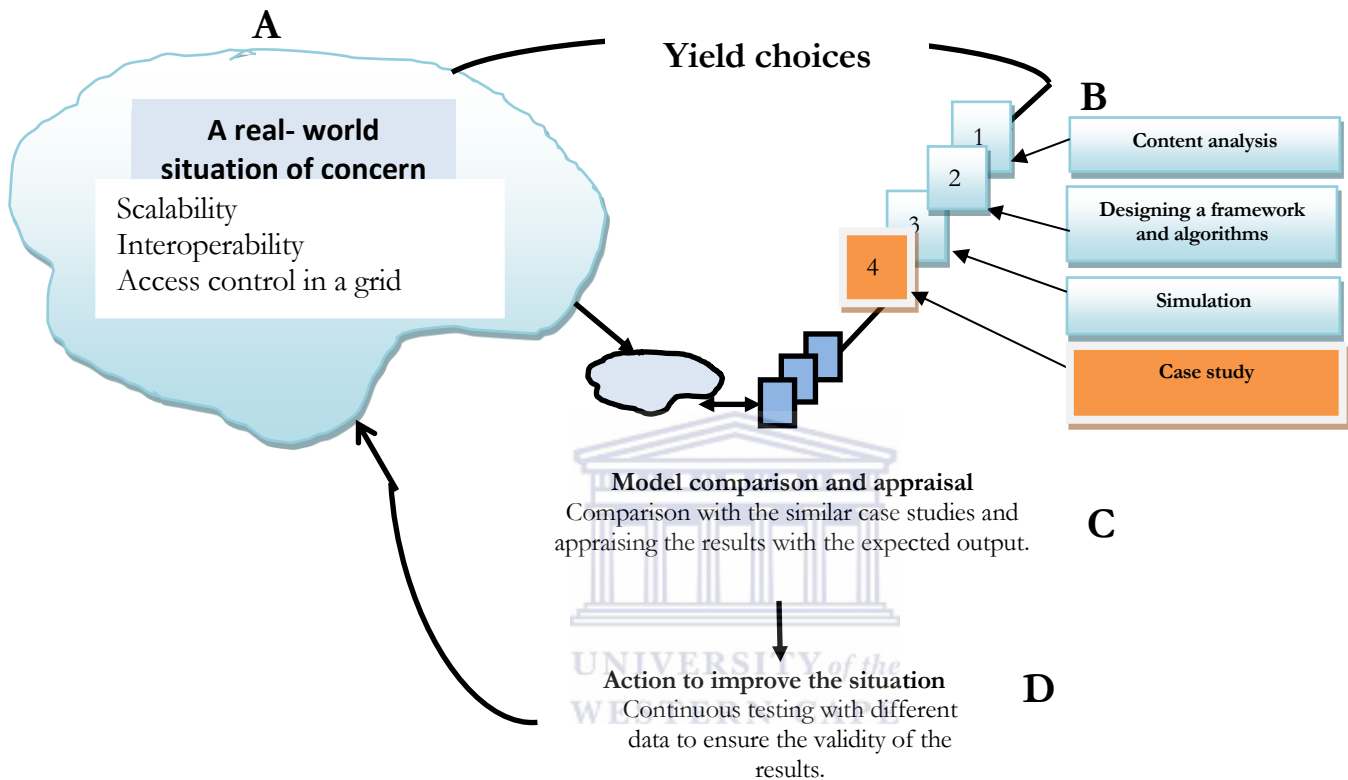


Figure 38: A case study approach using SSM

The last stage of SSM application is the case study to analyse and implement a real word situation of concern (i.e. access control) as depicted in Figure 38 .

#### SSM cycle 4: Case study

##### *A: Situation of concern*

Efficient access control has been identified as the challenge. It is the last stage of implementation being considered in this work.



*B: Alternative way of solution*

The RBAC model introduces role hierarchies (RH). Role hierarchies are defined for positioning and structuring various specified roles to reflect and present an organisation's line of responsibility and authority (Saeed & Reena, 2012 ). With role hierarchy, there is a role/role relation which defines user membership along with their various privilege inheritances. It presents organisational functional delineations and structures.

A typical role hierarchy as implemented for 3DGBE is presented in Figure 39. By standard and ab initio convention of role hierarchy, more strong and powerful (senior) roles are placed over those at the lower level (junior). As for 3DGBE, hierarchies are specified for the three domains with H1, H2 and H3 for domains A, B and C respectively.

A case study of e-health is used for each of the three domains. Different roles, permissions, domain hierarchies are specified. Using hierarchical RBAC, the Physician role which is defined in domain A with hierarchy H1 can have additional permissions to those from other two domains (Domains B and C). The inheritance of permission is transitive in nature across the three domains. For example, Physician, Cardiologist and Neurologist have different permissions already assigned to them in their domain (Domain A). They also have multiple permissions from other domain as specified in Table 7.

Also, Patient, Nurse, Pharmacist and Psychiatrist have their permissions domiciled in Domain B, the hierarchical nature of RBAC has their access permission extended to Domain C.

In short, Domain A roles with hierarchy H1 could access information in domain B and C. Domain B roles with hierarchy H2 has permissions to services in its domain as well as domain C. Finally, only Domain C role could not be permitted to access further information from any other domain except those defined within its hierarchy (H3).

If however, a GU in a domain whose hierarchy is lower decides to gain permission or intends to access resources in a higher hierarchical domain, such a request shall be denied.

Table 7: Multiple inheritance of permission in a 3DGBE

Role	Domains	Permission	Multiple inheritance of permission
Physician, Cardiologist, Neurologist, Obstetrician, Pathologist, Pulmonologist, Surgeon, Pediatrician, Oncologist, Dermatologist	A	P{A}	P{A}, P{B}, P{C}
Patient, Nurse, Pharmacist, Dentist, Psychiatrist, Podiatrist	B	P{B}	P{B}, P{C}
Ultrasound Technologist, X-Ray Technician, Clinical Technologist, Clinical Technologist, Dental Assistant, Dental Laboratory Technician	C	P{C}	P{C}

A case study scenario of an e-health system using the RBAC was implemented Hierarchical RBAC was adopted. Domains A, B and C were given hierarchies H1, H2 and H3, respectively.

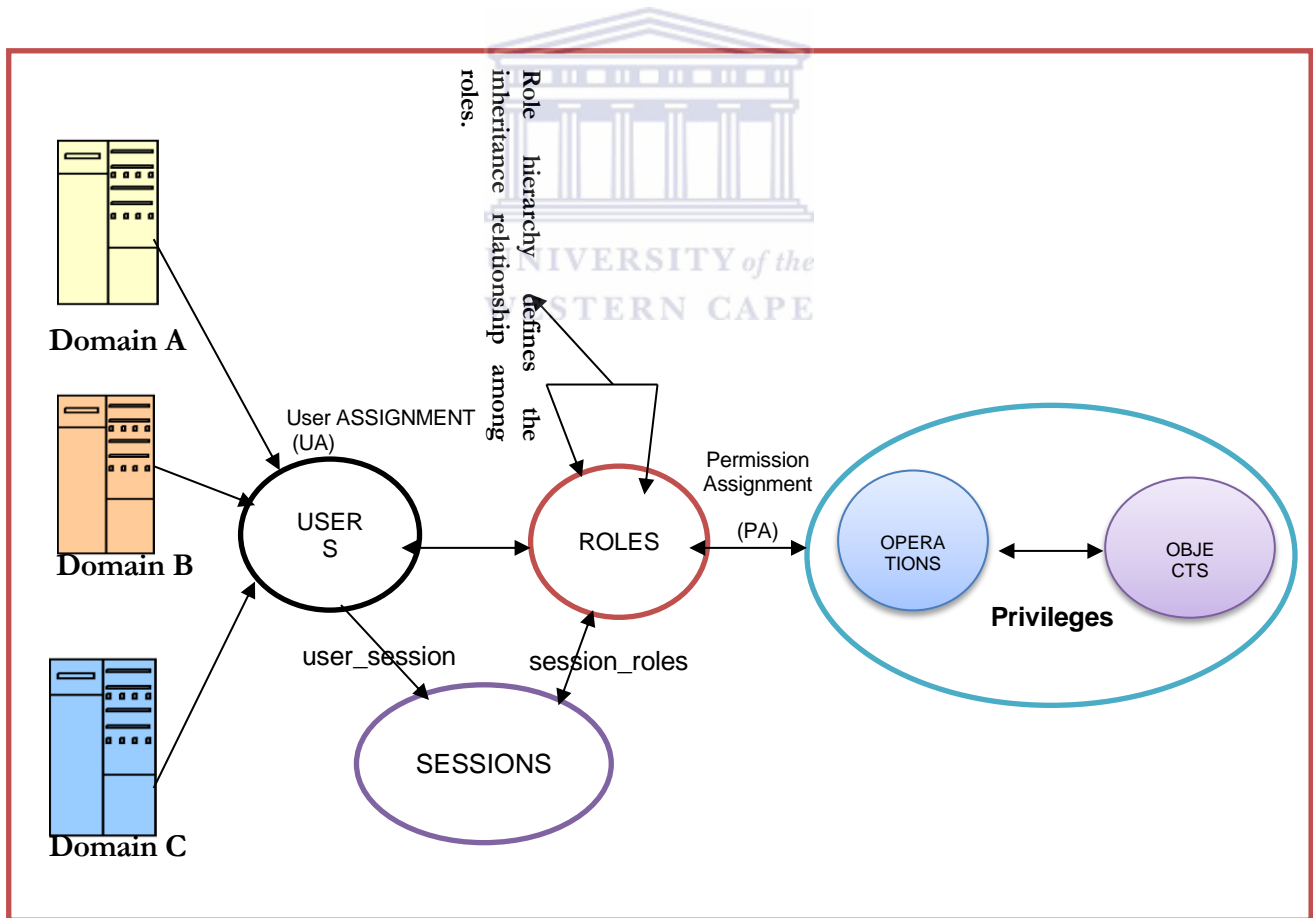


Figure 39: Implementation of hierarchical RBAC in 3DGBE

The implementation of access control was carried out successfully with the implementations of the architecture depicted in Figure 39 and Figure 40 respectively. Each domain has well defined roles for each of the grid users, depending on the hierarchy of his/her domain.

In Figure 40, the procedure through which hierarchical RBAC (Ren, Zuo, Li, Niu, & Yang, 2010) in 3DGBE was implemented is shown. A user coming from either Domain A, B or C has his/her identity authenticated by both the LSMU and CSMU respectively. The individual user's assigned role is monitored and guided by the corresponding domain hierarchy. The privileges available within the three domains are determined thereafter.

#### Terms and definitions as used in this context

- Let H1, H2 and H3 denote the hierarchies and let the role hierarchy (RH) denoted as H1, H2 and H3 be assigned to Domains A, B and C, respectively where  $H1 > H2 > H3$

It could be recalled that:

*“hierarchy is mathematically a partial order defining a seniority relation between roles, whereby the seniors' roles acquire the permission of their juniors, and junior roles acquire the user membership of their seniors”*

(Ferraiol, Sandhu, Gavrila, Kuhn, & Chandranouli, 2001).

- Let Role\_Domain A denote all roles defined in Domain A.
- Let Role\_Domain B denote all roles defined in Domain B.
- Let Role\_Domain C denote all roles defined in Domain C.
- **Resources**- can be any system resource that could be subject to access control, such as database record, a file, etc.
- **Operation** – could be an executable part of a program which, when invoked causes some basic function for the user to be executed.

- **Permission**-is an authorized approval given to an entity to carry out/perform operation on objects protected by RBAC access policy.
- **Role**- is the assigned function within the context of a particular domain of a grid network along with the associated semantics concerning the responsibility and authority given to a user to carry out.
- **User (grid user)** –is a person who requires access privilege to resources from one domain to another in a grid-based environment.

For implementation purposes therefore, hierarchical role-based access control was adopted to specify roles, services and permissions for each user from any domain. A scenario (Health) was considered in with each of the domains had roles, services and permission defined among the users. E-health was used as a case study because of the sensitive nature of the services, roles and permissions given to people involved in health-related services.

*Algorithm 3: Algorithm for efficient access control in a 3DGBE*

---

Required: Domains A, B and C, LSMU, CSMU

Grid user (GU) identification;

Get the Domain's hierarchy as {H1, H2, H3};

Assign hierarchy to the chosen domain;

Obtain GU role;

Retrieve GU services - permission;

Proceed to the grid

---

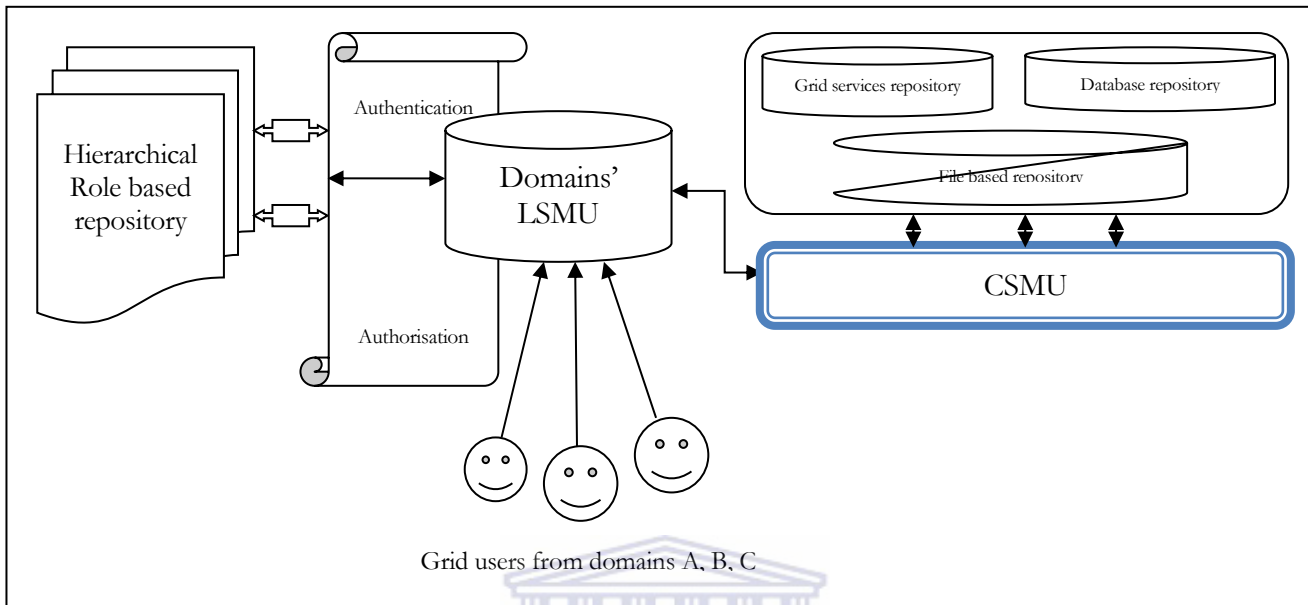


Figure 40: Description of hierarchical RBAC access control policy for 3DGBE

The full-scale description of how hierarchical RBAC was implemented for 3DGBE is shown in Figure 40. A GU from either Domain A, B and C is first subjected to authorisation and authentication by their respective LSMU and CSMU and thereafter have their domain hierarchies confirmed.

After this has been achieved, the final security checking of GU by the CSMU will be carried out, where each service requester has his/her roles, services and access privileges retrieved from the “database repository” and “grid services repository” respectively.

For the purpose of clarity, roles, services and permissions description for Domain A, B, and C are presented below.

#### *Role and services specification for DOMAIN A*

- **Role\_Domain A** = {Physician, Cardiologist, Neurologist, Obstetrician, Pathologist, Pulmonologist, Surgeon, Pediatrician, Oncologist, Dermatologist}
- **Services (permission)**

- i. {Physician (*write patient record, read patient record, write prescription, read prescription, examine patient*)}
- ii. {Cardiologist (*treat heart disease, write patient record, read patient record, write prescription, read prescription*)}
- iii. {Neurologist (*treat brain, examine nervous system, write patient record, read patient record*)}
- iv. {Obstetrician (*provide pregnancy care, delivers babies, write patient record, read patient record*)}
- v. {Pathologist (*interpret disease, examine tissue, read patient record, write prescription, read prescription*)}
- vi. {Pulmonologist (*treat respiratory problems, write prescription, read prescription*)}
- vii. {Surgeon (*perform operations, write patient record, read patient record*)}
- viii. {Pediatrician (*treat children, write prescription, read prescription*)}
- ix. {Oncologist (*treat cancers, read patient record*)}
- x. {Dermatologist (*treat skin, write prescription*)}

*Role and services specification for DOMAIN B*

- **Role\_Domain B** = {Patient, Nurse, Pharmacist, Dentist, Psychiatrist, Podiatrist}
- **Services (permission)**
  - i. {Patient (*read prescription, read patient record*)}
  - ii. {Nurse (*write patient record, read prescription, read patient record*)}
  - iii. {Pharmacist (*read prescription, read patient record, select prescription*)}
  - iv. {Dentist (*treat teeth, carry out operation, read prescription, read patient record*)}
  - v. {Psychiatrist (*treat mental patient, recommend drug, read prescription, read patient record*)}
  - vi. {Podiatrist (*treat human foot, treat lower leg, treat heel spurs, read patient record, write prescription*)}

*Role and services specification for DOMAIN C*

- **Role\_Domain C** = {Ultrasound Technologist, X-Ray Technician, Clinical Technologist, Clinical Technologist, Dental Assistant, Dental Laboratory Technician}
- **Services (permission)**
  - i. {Ultrasound Technologist (*read patient record, take ultrasound, analyse images*)}
  - ii. {X-Ray Technician (*read patient record, perform x-ray on patient, interpret and analyse x-ray result*)}
  - iii. {Clinical Technologist (*read patient record, perform medical test, interpret result*)}
  - iv. {Dental Assistant (*assist in dental operation, clean decayed teeth, give drug to patient*)}
  - v. {Dental Laboratory Technician (*read patient record, performs lab analysis*)}

A typical relationship indicating what is obtainable in a 3DGBE for GU, roles and permissions relationship is presented in Figure 41.

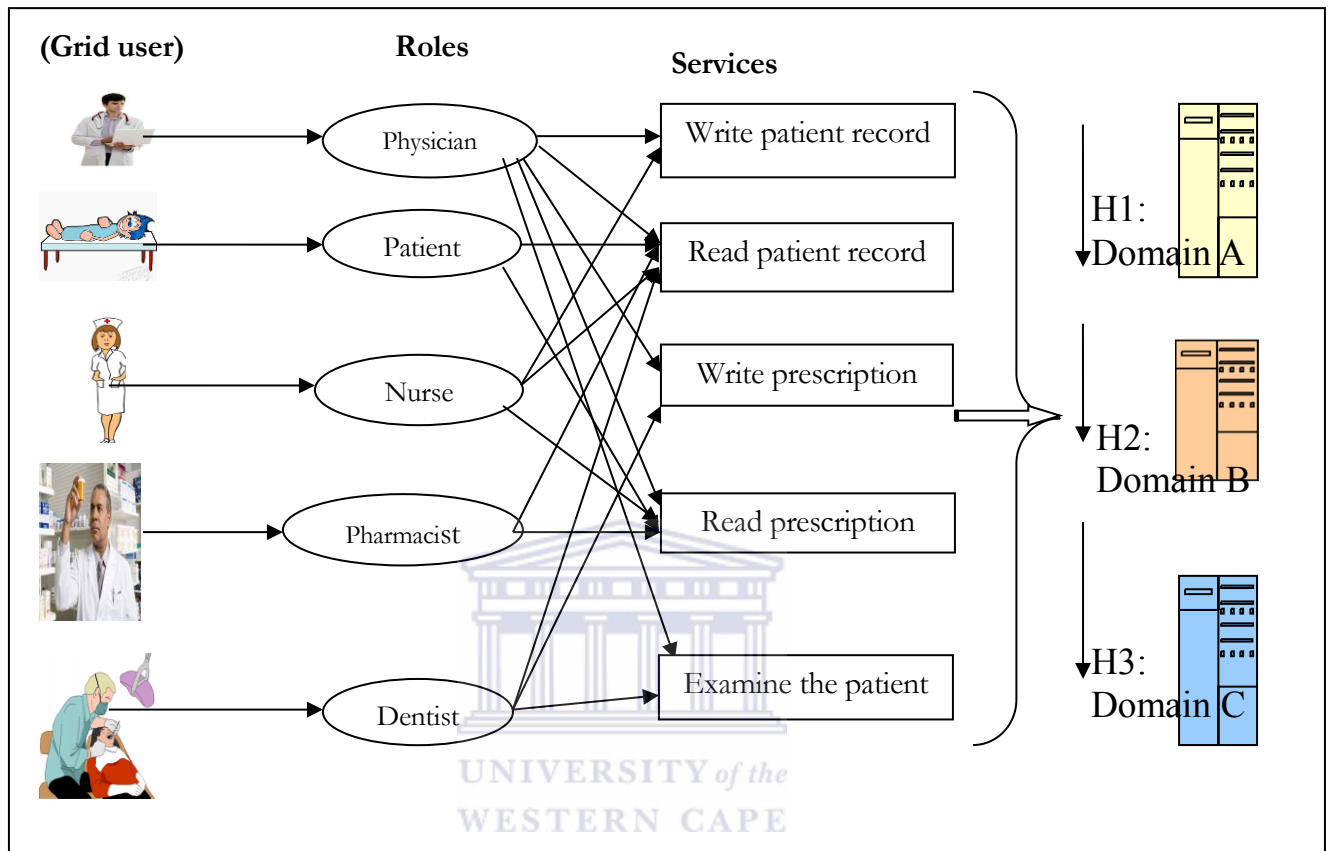


Figure 41: Grid user, roles and services relationship of hierarchical RBAC for 3DGBE

Whenever a GU specifies his/her domain, the corresponding hierarchy of such a user will be instantly verified and produced. The hierarchy is divided into three layers; Hierarchy 1 (H1) for domain A, Hierarchy 2 (H2) for Domain B and Hierarchy 3 (H3) for Domain C.

Of all the three domains' hierarchies, H1 has the highest rating in terms of accessibility to all available resources on the grid. Any GU who possesses H1 as his/ her own hierarchy implies he is from Domain A and has a capability of accessing resources from any desired domain whose services are defined.

The formulation is such that  $H1 > H2 > H3$ . This connotes the superiority of H1 over H2 and H3. Also, H2 has superiority over H3. With implementation, it implies that H1, with



the highest hierarchy, could access all the resources within his/her domain and the domains under it, i.e. Domains B and C, with H2 and H3, respectively.

Similarly, H2 could permit a GU with that hierarchy to access all available information in its domain and in the one below it, i.e. in H3. However, H3 permits the GU only to access resources within his/her domain alone. This initial access control framework is efficient in a 3DGBE, as users whose identities are not linked to a specific hierarchy will automatically be denied access to resources.

The prototype of this algorithm was coded and run in a Java Runtime Environment 1.7.0.5 for the workflows that define the model's task. A database consisting of the table and queries were created. Tables were created for domains, permissions, roles, services and users, while the queries were issued on the accessible resources and grid user. Implementation reveals that the hierarchical access control adopted for this scenario is reliable, dependable and efficient as an access control mechanism within the three domains considered.

Figure 48 (see Appendix D) gives an interface that pops up whenever a GU is successfully authenticated as a user in any of the three domains. These requirements have to be fulfilled by the GU in order to obtain efficient access to various resources when coming from his/her domain.

In a bid to track the GUs with their respective defined roles as specified and available on the grid, Figure 49 (see Appendix D) gives comprehensive information about domain roles and their services as spelt out for each user. From the foregoing, it is clear that a cardiologist who has his roles defined in Domain A of H1 has the corresponding listed services allocated to him.

Also, a dentist whose domain is B with domain hierarchy H2 can only access the allotted services. Attempt to access any other information or services from a domain different from

his will result to “rejection or denial of service”, which signifies the efficiency of the access control that is put in place. Finally, the same condition is applicable to the ultrasound technologist, who has his services defined in Domain C with hierarchy 3 (H3).

Figure 50 (see Appendix D) gives details about the status of grid users across different domains. Some grid users have their requests terminated because, the hierarchies and domains specified are not in order with information available from their domains. Those whose remarks and status are validated and proceed are given access to authorized information concerning them with respect to the domain and its hierarchy.

As can be seen from Figure 50 (see Appendix D), Saheed Dames, who specified his domain as A with hierarchy 1, has his access request terminated simply because he does not belong to the domain he specified. The same condition holds for Felix Adams and Kane Peter, among others.

Permission were granted to Kola Serifa (a surgeon) from Domain B , Andy Lui (an oncologist) and Nurayhn Wale (a podiatrist) because they specified domains where their roles are domiciled.

To further strengthen the accessibility to different resources and to maintain accuracy and precision in information concerning a user, Figure 51 (see Appendix D) gives a summary of services belonging to a user (Azeez, Abdula) whose role and services are domiciled in Domain B. He can only view this because he has been authorized and authenticated as a genuine and authorised user.

With the access control in place the GUs' access to various resources are well monitored, and controlled, and efficiently coordinated. This access control allows authorised access to resources across the three domains available on the grid. It easily provides avenue for screening out unauthorised users who are not supposed to access a particular resource and

permits legal and authorised user to access resources based on their defined roles and permissions from their respective hierarchical domains.

*C: Comparison and appraisal*

We identified the links and relationship between problem (access control), intervention (solution offered) and outcome (result) and we finally compared the differences, including cases selected for their different characteristics.

*D: How to improve the situation*

Repetitive testing was carried out to ensure that the results obtained conform with the expected outcome.

## CONCLUSION

In this chapter, the results obtained for solving the three different grid computing challenges (scalability, interoperability and efficient access control in a 3DGBE) were discussed. It could be recalled that the research questions set out at the beginning of this thesis focus at addressing, the challenges emanating from scalability, interoperability and efficient access control in a triple-domain grid-based environment, it is therefore important to examine whether these questions have been addressed using the methodology adopted.

The result obtained via content analysis was presented; this gives room for the summarized Table 4 which provides the features of the five security models discussed.

The second stage of the SSM cycle, witness the presentation of the architecture which comprises the phases involved in the designed architecture, the stages of the architecture, the algorithm that describes the working relation of the components in the architecture presented in stage 1. Also, the stage 2 of 3DGBE architectural framework was presented along with the algorithm.

Consequently, this chapter has fundamentally provided the various results obtained. Some of the results obtained in terms of scalability are presented. At the initial stage of evaluation, the new architecture (3DGBE) was compared with the existing MAC access control and the

result obtained proves the former is more scalable compared to the latter. Other parameters such as “average turnaround time” and “the number of grid nodes”; “average turnaround time” and “number of grid service requesters” were evaluated and finally the effect of “number of grid nodes” was also measured against “throughput” and the “grid nodes”.

Also, in terms of interoperability, the results have been presented using the four approaches adopted in this thesis for addressing this challenge. Interoperability was achieved through middleware integration across the three domains as shown in Figure 33. Apart from this, interoperability across 3DGBE was attained through the appropriate handling of authentication and authorisation between the LSMU and CSMU as presented in stage 1 of the architecture that is presented in Figure 22.

Also, results were presented in terms of the operating system interoperability and federated database approach to support the interoperability nature of the architecture.

This chapter has also presented a comparative evaluation of 3DGBE with other four basic access control models. The summary of the result is presented in Table 6.

Finally, the chapter present, using a case scenario, the application of hierarchical RBAC in 3DGBE. For each of the domains, independent roles and permissions are specified and their corresponding hierarchies are also stated. The results to support its evaluation are obtained and presented in the Appendix.

Based on the research question posed, it is therefore apparent that the results obtained have answered the questions raised.

*Chapter 5*

FINDINGS: INTERPRETATION OF THE RESULTS

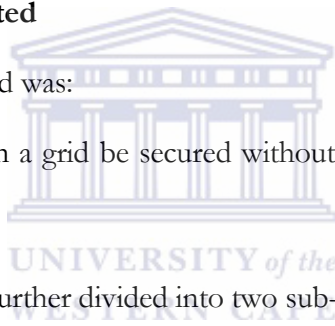
**INTRODUCTION**

Chapter 4 presents results obtained via the application of SSM. In this chapter, various findings from each cycle of the SSM, as well as those obtained from the previously stated research questions, will be interpreted.

**Research questions revisited**

The main question addressed was:

How should information on a grid be secured without compromising the accessibility and availability of resources?



The research question was further divided into two sub-questions as follows:

- How should interoperability and scalability be put in place to ensure optimal utilization of grid computing?
- How should efficient access control be administered for the entities on the grid in order to monitor and control permission to access various resources?

In order to adequately interpret the results obtained in terms of the research questions posed, the relationship among the following items needs to be established: the questions posed, the findings, the solution adopted, the effect of causal relationships and outcome indicator.

### **How should interoperability and scalability be put in place to ensure optimal utilization of grid computing?**

This question is split in two to address interoperability and scalability independently.

#### *a. Interoperability*

Findings from the existing literature revealed that the challenges of interoperability were addressed using different approaches such policy aggregation, trust, web ontology, modular information provider and public key infrastructures. In an attempt to finding solution to this challenge, the following strategies were explored in a 3DGBE: middleware aggregation, operating system and authentication and authorization between an LSMU and a CSMU, as well as federated database approach with MQL as a cross-domain query. Parameters identified and used for evaluation ranges from the gLite, Globus and Alchemi for middleware aggregation; UNIX, Linux and Windows (operating systems); and different applications across the three domains. Dynamic changes enforced on these parameters assisted in the efficient measurement and determination of interoperability. The result obtained showed the causal relationships provide the desired result as set out in the research question.

#### *b. Scalability*

A communication model, a framework based on a uniform abstraction and an authorisation framework are a few of the approaches used in addressing the concept of scalability in the existing literature.

Various metrics such as the number of accessed resources, time taken, average turnaround time, the number of grid node, the number of grid users, the number of service requesters, throughput and the number of nodes of the 3DGBE architectural model and algorithms were considered and measured against each other. The variation of the metrics used during the simulation process of the architecture ensured the verification of scalability. Various results obtained proved that the causal relationship effect significantly provide the desired result of scalability as set out in the research question. Various results as presented in Chapter 4 substantiate this argument.

*c. Access control*

**How should efficient access control be administered for the entities on the grid in order to monitor and control permission to access various resources?**

Semantic web ontologies and traditional access control models such as MAC, DAC, ACL, DA and CBSM were found in the existing literatures for controlling access in a grid-based environment. The flexible, dynamic and hierarchical RBAC was adopted. In Chapter 4, it was clearly shown how this model was used in specifying role, permission and services in 3DGBE. The implementation of flexible and hierarchical RBAC within the three domains of the developed architecture shows that effective access control is obtainable and achievable using this access control model. The results obtained prove the effectiveness of the design.

**Review of research goals.**

It is worth reminding the reader at this stage of the reason why this research was undertaken.

Having noted scalability, interoperability and access control as great challenges to the security of a grid computing system, the research aimed at addressing these challenges. This was achieved by first investigating various models relating to the challenges and developing a unique model based on the weaknesses noticed in the existing ones.

*Findings of the content analysis with SSM approach.*

To avoid subjectivity (Matveev, 2002) in identifying scalability, interoperability and access control as the security challenges to be addressed, this approach uses the repetitive approach in clarifying, ascertaining and pinpointing these challenges. The method assisted in discovering that 80% of reviewed papers stated that scalability is a serious security challenge in grid computing. Findings also revealed that 93% and 87% of the reviewed papers identified interoperability and access control respectively, as serious threats to the optimal utilization of a grid-based computing.

*Findings of the designed framework and algorithms.*

Since the objective of this research is to design an architectural framework to address these challenges, SSM was also used to design a workable, efficient and dependable framework for addressing them.

Five different models were initially designed before eventually arrived at the last design. This is the beauty of SSM, as it allows for a cyclic approach in addressing each of the challenges sequentially. Previous designs were redesigned to circumvent various challenges noted in the existing models.

The final design was compared several times with the existing models and its performance evaluation reflects the efficiency and effectiveness of the design.

*Findings of the simulation*

Findings show that SSM is undoubtedly beneficial in carrying out a valid empirical evaluation of the design. In order to obtain valid simulation results, we made several comparisons in terms of what is to be achieved (target results) and what has been achieved in practice. Efforts were also made to vary the causal relationship between various metrics adopted for simulation. Findings reveal that as the metrics were compared a number of times, more valid results were obtained. It is noted that if the simulation and experimentation were not tested several times, consistent, valid results may not have been achieved. Of 98% repeated simulated results, 91% remained consistent, which proves the evaluation results to be valid.

**Motivation for adopting SSM research methodology**

SSM methodology was considered suitable for this research based on the reasons listed below:

- i. It provides an explicit conceptual framework to deal with some of the grid security challenges observed.



- ii. It allows the carrying out of a cyclical-testing approach at different level in order to quickly observe the efficiency of the approach adopted.
- iii. It provides a mean of carrying out simulation until a desirable solution is attained.
- iv. It allows one to understand and identify loopholes that may arise from the solution. This results from comparing the expected solution with the solution arrived at.
- v. SSM provides opportunity to improve the problem design when the initial solution is not effective;
- vi. It provides a comparative analysis of the result obtained in relation to the target solution. Hence, it allows one to decide attempt to improve on the obtained result or not.

## CONCLUSION

In this chapter, various interpretations of results have been provided. In the first cycle of the SSM, a successful review of related literature was undertaken and the challenges addressed in this work were identified. The second stage of the cycle assisted in designing an efficient architectural framework and algorithms to handle the challenges identified in the first cycle. The cyclical nature of this approach helped to correct and redesign the model before arriving at the final framework. The unique feature of the framework lies in its ability to repeatedly carry out simulation and various types of experimentation in order to obtain dependable and reliable results.

The case study used in the experiment produced impressive results using this same approach. The cyclical nature of SSM has not only yielded good results, but has provided confirmed the validity of the results obtained through repetition of the process of experimentation and simulation. The next chapter will present the conclusion and recommendations.

*Chapter 6*

## DISCUSSION AND CONCLUSION

In the previous chapter, the research results were interpreted. This chapter evaluates the contribution of this thesis to the body of knowledge on distributed systems specifically grid computing. Also, conclusions are drawn based on the summary of what has been covered in the research as presented in the preceding chapters. Recommendations are then made for further research.

Grid computing technologies symbolize a significant and remarkable achievement towards the integration as well as aggregation of various network resources for solving complex, compute-intensive as well as very large-scale data-intensive applications. These features could be achieved at any location whenever a grid user intends to use grid resources.

Having noted that scalability, interoperability and efficient access control stand as stumbling-blocks for the full scale optimization and utilization of this form of distributed system, this thesis proposes some novel approaches for handling these security challenges. Specifically, a novel architecture and algorithms were developed and a case scenario was used. The architecture was evaluated using the Grid Security Services Simulator (which was released to the department from France for the purpose of this project), a meta-query language and Java Runtime Environment 1.7.0.5 for implementing the workflows that define the model's task. In this chapter, further explanation is given on the thesis contributions as well as possible future directions in this area of study.

**The need for research contribution**

This thesis simultaneously describes and solves the three basic security challenges in a grid-based environment. As evident from various citations provided in this thesis, each of these challenges have been independently addressed in various researches but not research has ever addressed them collectively as provided in this thesis.

This thesis presents a novel architecture christened as triple-domain grid-based environment (3DGBE). This unique architecture is divided into two different stages with each stage has two phases. The phases of the architecture are presented Figure 21 while both stage 1 and 2 of the architecture are presented in Figure 22 and 25 respectively. The choice of a 3DGBE is basically to create and establish competition among the various entities (users and resources) on the grid.

Furthermore, this thesis presents algorithms for both stages 1 and 2 (see algorithms 1 and 2) which provide comprehensive working relation of components of the architecture of Figure 22 as well as the working scenario of the architecture in Figure 25. The architecture has been successfully implemented using G3S.

This thesis provides four approaches for solving challenges emanating from interoperability. This approach of solving interoperability in a 3DGBE could be regarded as four-in-one approach. This is another unique contribution that needs emphasis. The thesis adopts grid middleware integration across the three domains as shown in Figure 33.

This approach is novel provides a successful interoperability across the three domains established. This thesis clearly presents how appropriate handling of authentication and authorization between the LSMU and CSMU in a 3DGBE has yielded interoperability. It is evident from the explanation on pages 80-81 how this was achieved. The third approach used in this category is operating system interoperability whose explanation is detailed on pages 81-83. Finally, different database applications were created across the three domains and meta-queries were issued in order to obtain desirable information across the three domain.

Another unique contribution of this research work to the body of knowledge lies in its adoption of hierarchical RBAC for providing efficient access control across the three domains. Since various grid users are involved and each of the GU could access resources anywhere, this this provides hierarchies for each user based on his domain and this was used as key for accessing resources. Each of the domains has a set of defined roles, permissions and their specified hierarchies. This aspect was evaluated using a case scenario of e-health.

The research presents a critical and extensive evaluation of five basic security models vis-à-vis their mode of implementations and operations, their strengths and their weaknesses. This evaluation has provided a clear-cut as to the best of the models suitable for the project. This thesis also presents unique approach for determining security rate (SR) across the three domain considered by taking into consideration “*Approv*”, “*Rep*” and “*Asses*” for all the entities in the 3DGBE.

All in all, the project is unique in simultaneously addressing three security challenges by using Grid Security Services Simulator (G3S), Java Runtime Environment 1.7.0.5 and a meta-query language.

Based on the contribution stated above, the need to further explain various applications of these challenges to aspect of human is indeed important.

A well-researched project, particularly a doctoral thesis, should provide a reasonable level of research contribution by supporting the advancement and enhancement of human knowledge by adding something unique, novel and original to the body of knowledge (Stella, 1984).

This can be attained by critically observing what problem(s) are currently being addressed and what type of results have been obtained after investigating these problems. Empirical and judgmental observation may spur research and identify new problems that are different from previous research (Baniassad & Clarke, 2004).

*According to Webster's Dictionary, knowledge is "the ideas or understandings which an entity possesses that are used to take effective action to achieve the entity's goal(s). This knowledge is specific to the entity which created it." (Steve, 2009:23)*

For knowledge to be established, new theories need to be formulated or existing ones manipulated. The development of a new theory often results from research question being posed to address gaps existing in the reviewed literature (Sanga & Venter, 2010).

Table 8: Summary of research contributions

Category	Typical Activity	Remarks
<b>Problem identification</b>	Identified three grid security problems, i.e. (scalability, interoperability and access control)	Achieved through SSM (review of conference and journal papers)
<b>Theoretical analysis</b>	Provided novel mathematical definitions and assumptions to aid simulation	Definitions of DSRV, SRV, Rep etc
<b>Design</b>	Designed a novel architectural framework for 3DGBE with algorithms	Tested and evaluated
<b>Comparison</b>	Compared several theoretical models, system designs, or implementations in a novel way	Achieved
<b>Implementation</b>	Implemented the designed system for the first time using G3S, MQL as “cross-domain” queries and Java Runtime Environment 1.7.0.5	Achieved
<b>Empirical analysis</b>	Studied the performance of an implemented system in a novel way by observing and comparing results with the stated objectives and research questions.	Effective and efficient
<b>Application of the project</b>	It has significance and numerous industrial and academic applications.	SAgrid, CHPC, Search for Extraterrestrial Intelligence (SETI), Indiana Grid Folding@home (Stanford University's chemistry department) and the military

The research identified various gaps in the literature. In this work, three basic security challenges were identified as forces militating against the full optimisation of grid computing.

Empiricism, novelty, acceptability to the research community (Balsley, 1970), disciplinary relevance and generalizability are considered as ingredients for regarding and classifying research as contributing to a body of knowledge (Burrell & Toyama, 2009).

Grid computing is a new area of research and, this research project is the first such project carried out in this area of computing in the department at the University of the Western Cape. The approach used is purely that of experimental computer science, which has to do with building and developing architectural framework and algorithms. This project has led to a couple of publications in leading computer science journals and presentations at Department of Higher Education and Training approved conferences.

Some of the research work have been published in peer-reviewed articles and presented at both national and in international conferences. (See pages ix – x). To further prove the acceptability of this research work, part of the work has been requested by Prof. Von Solms Rossouw of Nelson Mandela Metropolitan University who is the editor of the South African Institute for Electrical Engineers for further review and possible consideration for publication. The research contributions are summarized in Table 8.

### **Scope and limitations of the research**

The research focused on three basic security challenges that hampered the full utilisation of grid computing. The implementation of the designed framework was divided into three areas: implementations for scalability, interoperability and efficient access control. The results obtained in this project were achieved using three different approaches. The results for scalability were obtained with the use of a simulator (grid security services simulator), those on interoperability were arrived at using Java Runtime Environment 1.7.0.5, meta-query language (MQL) for “cross-domain” queries and experimentation, while the third phase (access control) of the results were achieved with the use of a case scenario using Java Runtime Environment 1.7.0.5 for its implementation. The results obtained so far have convincingly answered the research questions posed in Chapter 1 of the thesis.

As stated earlier, many challenges have faced grid computing since it came into existence. Three of them have been successfully addressed in this work. In other words, other challenges affecting grid computing such as grid scheduling, pricing, grid management, grid coordination and many others are beyond the scope of this work, hence they are not addressed in this work but deserve urgent attention.

The cyclical nature of the simulation and various experiments conducted is time consuming, coupled with the fact that it is a nascent area of research endeavour. All in all, the objectives of the project have been achieved, although, there is a need for implementation on a real-life grid.

### **Recommendations and future work**

It is evident from the thesis that all the challenges addressed are of essence and cannot be ignored either by security architects or security experts. However, since the architectural models were not implemented in a real grid environment, it therefore advisable to do so in order to test their applicability in the real world.

This thesis suggests some future directions to further solve other challenges hampering the full-scale adoption and utilization of grid computing. The future directions are related to the following key functionalities of grids, i.e. resource management, data management, scheduling, resource integration, coordination, pricing and auditing, implementation on a real grid testbeds, and extension of implementation to other forms of distributed systems.

Another interesting area of grid that deserves urgent attention is negotiation and coordination of multiple resources by various organisations. For example, in the case of conserving and reserving network bandwidth, a network engineer is required to emphasis and focus on network management issues like providing and establishing a dependable and guaranteed handling traffic congestion mechanism and end-to-end path for resources.

The study about scalability, interoperability and efficient access control were carried out through simulation, java implementations, case scenario and meta-query language. The next

stage is to convert this research study into a real life implementation, i.e. by implementing them as a prototype on a real grid testbeds.

Apart from grid computing, there are other forms of distributed system that are also facing challenges addressed in this thesis. Specifically, cloud computing, peer to peer system and cluster computing also deserve attention while implementing these challenges.

Finally, there is a strong believe that real life implementation of these three challenges will go a long at addressing industrial and academe problems being faced through them.





## BIBLIOGRAPHY

- Abbes, H., Barbera, R., Jemn, M., & Mazigh, S. (2008). Towards the Interoperability between GTRS and EUMEDGRID. *2008 IEEE Asia-Pacific Services Computing Conference* (pp. 674-679). IEEE Computer Society.
- Akimana, R., & Markowitch, O. (2006). Data and Code Integrity in Grid Environments. *Proceedings of the 6th WSEAS International Conference on Simulation, Modelling and Optimization*, (pp. 677-682). Lisbon, Portugal.
- Al-Bayatt, A., Zedan, H., & Siewe, F. (2009, February). Access Control Mechanism for Mobile Ad hoc Network of Networks (MANoN). Software Technology Research Laboratory, De Montfort University, Leicester, UK.
- Alfawair, M., Aldabbas, O., Bartels, P., & Zedan, H. (2007). *Grid Evolution. International Conference on Computer Engineering & Systems (ICCES'07)* (pp. 158-163). Cairo , Egypt : IEEE Computer Society.
- Anderson, R. (2008). *Security Engineering: a guide to rebuilding dependable distributed systems*. New York, USA: Wiley, Computer Publishing.
- Ashley, P., & Karjoth, G. (2003). Shortcomings of P3P for Privacy Authorization. *W3C Workshop on the long term Future of P3P and Enterprise Privacy Languages* (pp. 1-13). Schleswig-Holstein, Germany: World Wide Web Consortium.
- Azeez, N. A., Iyamu, T., & Venter, I. M. (2011). Grid Security Loopholes with proposed countermeasures. In E. Gelenbe (Ed.), *ISCIS2011, 26th International Symposium on Computer and Information Sciences 26-28 September 2011* (pp. 411-418). Imperial College, London, UK: Springer Verlag.
- Azeez, N., & Osunade, O. (2009). Towards ameliorating cybercrime and cybersecurity. *(IJCSIS) International Journal of Computer Science and Information Security*, 3(1), 1-11.
- Azeez, N., & Osunade, O. (2009). Towards ameliorating cybercrime and cybersecurity. *(IJCSIS) International Journal of Computer Science and Information Security*, 3, No. 1,, 1-11.
- Aziz, B., Foley, S. N., Herbert , J., & Swart, G. (2006). Reconfiguring Role Based Access Control Policies Using Risk Semantics. *Journal of High Speed Networks*, 261-273.
- Baker, M., Buyya, R., & Laforenza, , D. ( 2002). Grids and Grid Technologies for Wide-Area Distributed Computing. *International Journal of Software: Practice and Experience (SPE)*,, 32, 15, 1437-1466.
- Baker, M., Buyya, R., & Laforenza, D. (2002). Grids and Grid Technologies for Wide-Area Distributed Computing. *International Journal of Software: Practice and Experience (SPE)*,, 32(15), 1437-1466.

- Baktash, H. A., Karimi, M., Meybodi, M. R., & Bouyer, A. (2010). 2L-RBACG: a New Framework for Resource Access Control in Grid Environments. *Fifth International Conference on Digital Information Management (ICDIM)* (pp. 359 - 366). Canada: IEEE Computer Society.
- Balsley, H. (1970). *Quantitative research methods for business and economics*. New York: Random House.
- Baniassad, E., & Clarke, S. (22 March 2004). Finding Aspects in Requirements with Theme/Doc. *proceedings of Early Aspects 2004: Aspect Oriented Requirements Engineering and Architecture Design*. Lancaster, UK.
- Barbera, R., Fargetta, M., & Giorgio, E. (2007). Multiple Middleware co-existence: another aspect of Grid Interoperation. *Third IEEE International Conference on e-Science and Grid Computing* (pp. 577-583). IEEE Computer Society.
- Berman, F., Fox, G., Hey, A. J., & Hey, T. (2005). *Grid computing: making the global infrastructure a reality*. West Sussex, England: John Wiley and Sons Ltd.
- Bogdan, R., & Taylor, S. J. (1998). *Introduction to qualitative research methods*. New York: John Wiley.
- Bouwman, B., Mauw, S., Eindhoven, P. R.-O., & Petkovic, M. (2008). Rights Management for Role-Based Access Control. *Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE* (pp. 1085 - 1090). Las Vegas, NV: IEEE Computer Society.
- Brown, M. D. (1999, April). *The International Grid (iGrid): Empowering Global Research Community Networking Using High Performance International Internet Services*. Retrieved from <http://www.globus.org/research/papers.html>
- Burrel, J., & Toyama, K. (2009,). What Constitutes Good ICTD Research? *Information Technology and International Development Journal*, 5(3), 82–94.
- Buttler, R., Welch, V., Engert, D., Foster, I., Tuecke, S., Volmer, J., et al. (2000, December). A National-Scale Authentication Infrastructure. *Nat. Center for Supercomput. Applications*, 33(12), 60-66.
- Buyya, R. (2002). PhD Thesis (Doctor of Philosophy). In Buyya, *Economic-based Distributed Resource Management and Scheduling for Grid Computing*. Melbourne, Australia: Monash University.
- Buyya, R. ( 2003). Global Data Intensive Grid Collaboration,HPC Challenge Proposal and Demonstration. *IEEE/ACM Supercomputing Conference (SC 2003)*.
- Buyya, R., & Venugopal, S. (2004). The Gridbus Toolkit for Service Oriented Grid and Utility Computing: An Overview and Status Report. *Proceedings of the First IEEE International Workshop on Grid Economics and Business Models (GECON 2004, April 23, 2004, (pp. 19-36)*. Seoul, Korea, : IEEE Press, New Jersey, USA.

- Buyya, Rajkumar. (2002). Economic-based Distributed Resource Management and Scheduling for Grid Computing. *Doctor of Philosophy*. Melbourne, Australia: Monash University.
- Buyya, Rajkumar. (2009, May 12). *Grid Computing Information Center (Grid Infoware)*. (the gridbus project) Retrieved March 19, 2011, from Grid Computing: <http://www.gridcomputing.com/>
- Buyya, R., Abramson, D., & Giddy, J. (2001). *A case study for economic grid architecture for service oriented grid computing*. Washington DC: IEEE Computing Society.
- Cai, Z., Liu, M., Guo, X., Zhang, Q., & Geng, F. (2009). A Password-based Authorization Management System Using AKE Protocol in Grid Systems. *International Conference on New Trends in Information and Service Science* (pp. 546- 551). Beijing, China: IEEE Computer Society.
- Cassell, C., & Symon, G. (1994). Qualitative research in work contexts. *Qualitative methods in organizational research*, 1-13.
- Chadwick, D., & Otenko, A. (2002). RBAC Policies in XML for X.509 Based Privilege Management. *SEC '02 Proceedings of the IFIP TC11 17th International Conference on Information Security: Visions and Perspectives*. The Netherlands, The Netherlands ©2002: ACM, Computing.
- Chadwick, D. (2005). *Authorisation in Grid computing*. University of Salford, UK: Elsevier.
- Checkland, P. (2000). *Soft Systems Methodology: A Thirty Year Retrospective*.
- Chen, J., Wu, H., Wang, Q., & Chi, X. (2006). A Reflective Framework for Authentication in Grid Computing Environments. *Proceedings of the Fifth International Conference on Grid and Cooperative Computing (GCC'06)* (pp. 404-407). Washington, DC, USA: IEEE Computer Society.
- Chen, Y., Luo, J., & Ni, X. (2008). A Fuzzy Trust Evaluation Based Access Control in Grid Environment. *The Third ChinaGrid Annual Conference* (pp. 190-196). China: IEEE Computer Society.
- Chetty, M., & Buyya, R. (2002, August). Weaving Computational Grids: How Analogous Are They with Electrical Grids? *IEEE Computer Society Press and American Institute of Physics, USA*, 61-71.
- Chon, R., Enokido, T., Wietrzsk, V., & Takizawa, W. (2004). Role Locks to Prevent Illegal Information Flow among Objects. *18th International Conference on Advanced Information Network and Application (AINA'04)*, 1, pp. 196-201. Fukuoka, Japan.
- Cirio, L., Cruz, I. F., & Tamassia, R. (2007). A Role and Attribute Based Access Control System Using Semantic Web Technologies. In V. Robert Meersman, T. Zahir, & H. Pilar (Ed.), *OTM'07 Proceedings of the 2007 OTM Confederated international conference on On the move to meaningful internet systems*. 2, pp. 1256-1266. Springer-Verlag Berlin, Heidelberg ©2007.

- Cline, A. (2011, October). *What is Epistemology?* Retrieved November 2, 2012, from Religion & Spirituality: <http://atheism.about.com/od/philosophybranches/p/Epistemology.htm>
- Crotty, M. J. (1998). *The foundations of social research: meaning and perspective in the research process* (Vol. 5). SAGE Publication.
- D' Arcy, D., & Hovav, A. (2007). Deterrign internal information system systems misuse. *Communications of the ACM*, 50(20), 113-117.
- D' Arcy, D., & Hovav, A. (2007). Deterrign internal information system systems misuse. *Communications of the ACM*, 50, no 20,, 113-117.
- Dallon, G., Massonet, P., Molderez, J. F., Ponsard, C., & Arenas, A. (2007). An analysis of the Chinese Wall Pattern for Guaranteeing Confidentiality in Grid-Based Virtual Organisations. *Security and P* (pp. 217-222). Privacy in Communications Networks and the Workshops, secureComm 2007: IEEE.
- Darke, P., Shanks, G., & Broadben, M. (1998). Successfully completing case study research: combining rigour, relevance and pragmatism. *Information System Journal*, 273-289.
- Detsch, A., Gasparly, L., Barcellos, M., & Cavalheiro, G. (2004). Towards a Flexible Security Framework for Peer-to-Peer based Grid Computing. *2nd Workshop on Middleware for Grid Computing* (pp. 52-56). Toronto, Canada: ACM.
- Dewan, P., Grundin, J., & Horvitz, E. (2007). Towards a mixed-initiative access control. *COLCOM '07: Proceedings of the 2007 International Conference on Collaborative Computing: Networking, Applications and Worksharing*, (pp. 64-71). New York, USA, November.
- Di, S., Jin, H., Li, S., Chen, L., Qi, L., & Wang, C. (2007). Ontology Based Grid Information Interoperation. *International Conference on Advanced Information Networking and Applications Workshops* (pp. 91-96). Ontario, Canada: IEEE Computer Society.
- Etalle, S., & Winsborough, W. (2007). A Posteriori Compliance Control. in *SACMAT '07: Proceedings of the 12th ACM symposium on Access control models and technologies* (pp. 11-20). Sophia Antipolis, France: IEEE Computer Society.
- Federico, S. (March 2009). on usage control for data Grid: models, architectures, and specification. In *PhD Thesis*. University of Ferrara.
- Ferreira, A., Cruz-Correia, R., Antunes, L., Farinnha, P., Oliveira-Palhares, E., Chawick, D. W., et al. (2006). How to break access control in a controlled manner. in *Proceedings of the 19th IEEE International Symposium on Computer-Based Medical System* (pp. 847-51). Salt Lake City, Utah: IEEE Computer Society.
- Foster, I., Kesselman, C., Tsudik, G., & Tuecke, S. (1998). A Security Architecture for Computational Grids. *ACM Conference on Computers and Security* (pp. 83-91.). New York, NY, USA: IEEE.

- Ferraiol, D. F., Sandhu, R., Gavrilu, S., Kuhn, D. R., & Chandranouli, R. (2001). Proposed NIST Standard for Role-Based Access Control. *ACM Transactions on Information and System Security*, 4(3), 224–274.
- Ferreira, L., Berstis, V., Armstrong, J., Kendzierski, M., Neukoetter, A., & Takagi, M. (2003). *Introduction to Grid Computing with Globus* (Second ed.). United States: International Business Machines Corporation 2002, 2003.
- Foster, I., Yong, Z., Raicu, I., & Lu, S. (2008). Cloud Computing and Grid Computing 360-Degree Compared. *Grid Computing Environments Workshop, 2008. GCE '08* (pp. 1- 10). Austin, Texas: IEEE Computer Society.
- Foster, I.; Kesselman, C.; Tuecke, S. (2001). The anatomy of the grid: Enabling scalable virtual organizations. *International Journal of High Performance Computing Applications*, 15, 200-222.
- Foster, Ian; Kesselman, Carl. (2003). *The grid2: Blueprint for a new computing infrastructure*. USA: Elsevier.
- Foster, Ian; Kesselman, C K. (1997). Globus: A Metacomputing Infrastructure Toolkit. *The International Journal of Supercomputer Applications and High Performance Computing*, 11(2), 115-128.
- Foster, I., Kesselman, C., Lee, C., Lindell, B., Nahrstedt, K., & Roy, A. (1999). A Distributed Resource Management Architecture that Supports Advance Reservations and Co-Allocation. *citeseerx*.
- Frankfort-Nachmias, C; Nachmias, D. (1992). *Research methods in the social sciences (4th ed.)*. New York: St. Martin's Press.
- Franz, B., Ian, H., & Ulrike, S. (2007). Description Logics. *Elsevier*, 1-47.
- Frey, J., Tannenbaum, T., Foster, I., & Livny, M. (2001). Condor-G: A Computation Management Agent for Multi-Institutional Grids. *Proceedings of the Tenth IEEE Symposium on High Performance Distributed Computing (HPDC10) San Francisco, California*, (pp. 34-40). California.
- Geethakumari, G., Atul, N., & Sastry, V. N. (2009). A Cross-Domain Role Mapping and Authorisation Framework for RBAC in Grid Systems. *IJCSA*.
- Gentzsch, W. (2000). DOT-COMing the GRID: Using Grids for Business. In B. Rajkumar, & B. Mark (Ed.), *Grid Computing-GRID 2000* (pp. 1-3). Bangalore, India: Springer.
- Georgiev, I. K., & Georgiev, I. I. (2001). A security model for distributed computing. *Journal of computing sciences in colleges*, 17 v1, 178-86.
- Gilbert, A., Abraham, A., & Paprzycki, M. (2004). A system for ensuring data integrity in grid environments . *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004* (pp. 435 - 439 ). OK, USA: IEEE Computer Society.

- Goguen, J., & Meseguer, J. (2009). Security Policies and Security Models. *SRI International* (pp. 1-10). Menlo Park, CA 94025: IEEE.
- Gor, K., Ra, D., Ali, S., Alves, L., Arurkar, N., Gupta, I., et al. (2005). Scalable Enterprise Level Workflow and Infrastructure Management in a Grid Computing Environment. *2005 IEEE International Symposium on Cluster Computing and the Grid. 2*, pp. 661- 667. Cardiff, United Kingdom: IEEE Computer Society.
- Gouglidis, A., & Mavridis, I. (2009). A Foundation for Defining Security Requirements in Grid Computing. *13th Panhellenic Conference on Informatics* (pp. 180-184). IEEE.
- Goyal, P. (2011). Application of a Distributed Security Method to End-2-End Services Security in Independent Heterogenous Cloud Computing Environment. *IEEE World Congress on Services* (pp. 379-384). Washinsgton, DC: IEEE Computer Society.
- Grid. (2010). *Welcome to South African Nation Grid Portal*. Retrieved November 22, 2010, from South African National Grid: <http://www.sagrid.ac.za/>
- Gridbus. (2009, May 23). *Grid Computing Info Centre (GRID Infoware)*. (R. Buyya, Editor) Retrieved August 1, 2010, from <http://www.gridcomputing.com/>
- Grimshaw, A. (2008, January 10). *A Worldwide Virtual Computer for an Advancing Legion of Applications, NPACI*. Retrieved April 3, 2012, from <http://www.npaci.edu/enVision/v15.2/legion.html>
- GStat. (2010, July 5). *GStat 2.0*. Retrieved August 2nd, 2010, from Summary View- Grid ALL: <http://gstat.gridops.org/gstat/sa-grid>
- Haidar, A. N. (2002, September 25). Critical Evaluation of Current Approaches to Grid Security. *MSc Degree project*. Royal Hooloway, University London, UK.
- He, Y., Li, F., & Hu, W. (2008). The design and implementation of security communication model in grid networks. *International conference in computer science and information technology* (pp. 421-424). China: IEEE Computer Society.
- Heimbigner, D., & McLeod, D. (1985). A federated architecture for information management. *ACM Transactions on Information Systems (TOIS)*, 3(3), 253 - 278.
- Herath, T., & Rao, H. R. (2009). Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, 18 n 2, 106-25.
- Hey, A., & Trefethen, , A. (2002, October). The UK e-Science Core Programme and the Grid. *Future Generation Computer Systems*, 18(8), 1017-1031.
- Higgins, G. E., Wilson, A. L., & Fell, B. D. (2005). An Application of Deterrence Theory to Software Piracy. *Journal of Criminal Justice and Popular Culture*, 12 no 3, 166-84.
- Houmb., S., Georg, G., France, R., & Matheson, D. (2004). Using aspects to manage security risks in risk-driven development. *3rd International Workshop on Critical Systems Development with UML*, (pp. 71-84). Lisbon, Portugal.

- Humphrey, M., Thompson, M., & Jackson, K. (2005). *Security for Grids*. Paper LBNL-54853. Lawrence Berkeley National Laboratory.
- Hu, V. C., Ferraiolo, D. F., & Kuhn, D. R. (2006). *Assessment of Access Control Systems*. National Institute of Standards and Technology. Gaithersburg: U.S Department of Commerce.
- Hwang, M.-S., & Yang, W.-P. (2010). A new dynamic access control scheme based on subject-object list. *Data & Knowledge Engineering*, 45-56.
- IASTED. (2004). *International Association of Science and Technology for Developmet*. www.iasted.com.
- Imine, A., Cherif, A., & Rusinowitch, M. (2009). *An Optimistic Mandatory Access Control Model for Distributed Collaborative Editors*. INRIA: IEEE.
- Jacob, K., & Fukui, B. (2005). Introduction to grid computing. In I. I. (Redbooks). IBM.
- Jeong, J., Yu, W., Shin, D., Shin, D., Moon, K., & Lee, J. (2006). Integration of Single Sign-On and Role-Based Access Control Profiles for Grid Computing. *Frontiers of WWW Research and Development - APWeb 2006*. 3841, pp. 880-885. Harbin, China: Springer Verlag.
- Junrang, L., Zhaohui, W., Jianhua, Y., & Mingwang, X. (2004). A secure for network-attached storage on the grid. *Proceedings of 2004 IEEE International Conference on service computing* (pp. 604- 608). IEEE Computer Science.
- Kaga, L., & Abelson, H. (1990, March). *Access Control is an Inadequate Framework for Privacy Protection*. Retrieved February 19, 2011, from www.dig.csail.mit.edu/2010/Talks/0712-W3CPrivacy-lk/privacy.pdf
- Kealey, D. J., & Protheroe, D. R. (1996). The effectiveness of cross-cultural training for expatriates: An assessment of the literature on the issue. *International Journal of Intercultural Relations*, 20(2), 141-165.
- Khider, H., Osman, T., & Sherkat, N. (2010). Attribute-Based Authorization for Grid Computing. *International Conference on Intelligent Systems, Modelling and Simulation*, (pp. 70-74).
- Kumari, A., Shakti Mishra, S., & Kushwaha, D. S. (2010). A New Collaborative Trust Enhanced Security Model For Distributed System. *International Journal of Computer Applications*, 18(26), 127-134.
- Laccetti, G., & Schmid, G. (2007 , June 5). A framework model for grid security. *Future Generation Computer Systems*, 23.
- Lakshmanan, L. V., Sadri, F., & Subramanian, S. N. (2001). SchemaSQL—An Extension to SQL for Multidatabase Interoperability. *ACM Transactions on Database Systems*, 26, 476-519.

- Lang, B., Foster, I., Siebenlist, F., Ananthakrishnan, R., & Freeman, T. (1990). *Attribute Based Access Control for Grid Computing*, (pp. 1-13).
- Levine, D. (2002.). Auditing Computer Security. In *Computer Security HandBook*. Fourth Edition, Wiley, .
- Li, J. (2005). A scalable authorization approach for grid system environments. In M. William , & W. S. Waleed (Ed.), *International Symposium on Collaborative Technologies and Systems* (pp. 84-91). SAINT Louis, MIssOURI, USA: IEEE Computing Society.
- Li, Y., Sun, H., Chen, Z., Ren, J., & Luo, H. (2008). Using Trust and Risk in Access Control for Grid Environment. *2008 International Conference on Security Technology* (pp. 13-16). Prague , Czech Republic: IEEE Computer Society.
- Lin, Q., Kang Neo, H., Liang, N., Guangbin, H. Z., & Gray, R. K. (2007). Grid-based large-scale Web3D collaborative virtual environment. *Web3D '07 Proceedings of the twelfth international conference on 3D web technology* (pp. 123-134). New York, NY, USA: ACM.
- Lupu, E., & Sloman, M. (1999, November). Conflicts in Policy-Based Distributed Systems Management . *IEEE Trans. Software Eng*, 852-869.
- Luther, A., Buyya, , R., Ranjan, , R., & Venugopal, S. (2005, March 24). "*Alchemi: A .NET-based Grid Computing Framework and its Integration into Global Grids*". Retrieved April 12, 2012, from <http://www.gridbus.org/papers/Alchemi.pdf>
- Lu, R., Cao, Z., Chai, Z., & Liang, X. (2008, September). A Simple User Authentication Scheme for Grid Computing. *International Journal of Network Security*, 7, 202–206.
- Machiraju, V., Sahai, A., & Moorsel, A. (2003). Managed Utility Computing: The Grid as Management Backplane. *Dependable in Computing: Lecture Notes in Computer Science*. 2847. Springer.
- Mao, Z., Li, N., Chen, H., & Jiang, X. (2009). Trojan horse resistant discretionary access control. in *SACMAT '09: Proceedings of the 14th ACM symposium on Access control models and technologies* (pp. 237-46). Stresa, Italy: IEEE Computer Society.
- Matveev, A. V. (2002). The Advantages of Employing Quantitative and Quantitative Methods in Intercultural Research. *Bulletin of Russian Communication Association "THEORY OF COMMUNICATION AND APPLIED COMMUNICATION"*, 59-67.
- McClure, S., Scambray, J., & Kurtz, G. (2003). *Hacking Exposed: Network Security secrets and solutions*. U.S.A: McGraw Hill: NY, U.S.A.
- McCroskey, J. C., & Richmond, V. P. (1990). Willingness to communicate: Differing cultural perspectives. *The Southern Communication Journal*, 56, 72-77.
- McLean, J. (2008). The Algebra of Security. *IEEE Symposium on Security and Privacy* (p. 2). Oakland, CA: IEEE Computer Society.



- Menezes, A. J., Vanstone, S. A., & Van Oorschot, P. C. (1997). *Handbook of Applied Cryptography*. Massachusetts Institute of Technology, USA: CRC Press, Inc.
- Mohteshim, H. (2005). Passive and Active attacks against Wireless LANS.
- MSDN. (2005, December). *Data Confidentiality*. Retrieved February 12, 2011, from <http://msdn.microsoft.com/en-us/library/ff650720.aspx>
- Naqvi, S., & Riguidel, M. (2005). Grid Security Services Simulator (G3S)- A Simulation Tool for the Design and Analysis of Grid Security Solution. *Proceedings of the First Inter'l Conference on e-Science and Grid Computing (e-Science '05)* (pp. 1-8). IEEE Computer Society.
- NHSE. (2009, April 5). Parallel Computing Research Newsletter. *National HPCC Software Exchange*, 5 (4), pp. 4-7.
- Nithya, M. W., & Banu, R. (2010). Towards Novel and Efficient Security Architecture for RBAC in Grid Computing. *Int'l Journal of Computer Science & Information Technologies*, 1(1), 16-23.
- NWICG. (2008, December 6). *Northwest Indiana Computational Grid (NWICG)*. Retrieved August 2, 2010, from Purdue University-Calumet, Purdue University – West Lafayette, and the University of Notre Dame: [www.nwicgrid.org](http://www.nwicgrid.org)
- Onieva, J. A., & Zhou, J. (2008). Secure Multi-Party Non-Repudiation Protocols and Applications. *Springer (Security and Cryptology)*, 43, 20.
- Pala, M., Cholia, S., Rea, S. A., & Smith, S. W. (2008). Extending PKI Interoperability in Computational Grids. In P. Thierry, L. Laurent, & B. Rajkumar (Ed.), *Eighth IEEE International Symposium on Cluster Computing and the Grid* (pp. 645-650). Lyon, France: IEEE Computer Society.
- Patrick, B., & Ghita, K. M. (2004). Context-Based Security Policies: A New Modeling Approach. *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops (PERCOMW'04)* (pp. 154-158). IEEE Computer Society.
- Pearlman, L., Welch, V., Foster, I., Kesselman, C., & Tuecke, S. (2010). A Community Authorization Service for Group Collaboration. 23-34.
- Pearlman, L., Kesselman, C., Welch, V., Foster, I., & Tuecke, S. (2003). The Community Authorization Service: Status and Future. *CHEP03*, (pp. 1-9). La Jolla, California.
- Pedayachee, K. (2009, December 2). An aspect-oriented approach towards enhancing Optimistic Access Control with Usage Control. *Doctor of Philosophy*. Pretoria, Northern, South Africa: University of Pretoria.
- Pinkas, B., & Sander, T. (2008, March). *Securing Passwords Against Dictionary Attacks*. Retrieved February 18, 2011, from [www.pinkas.net/PAPERS/pwdweb.pdf](http://www.pinkas.net/PAPERS/pwdweb.pdf)
- Pfleeger, C. (1997). *Security in Computing* (2nd ed.). Prentice Hall, United States of America: Englewood Cliffs.

- Pretschner, A., Hilty, M., Schutz, F., Schaefer, C., & Walter, T. (2008). Usage Control Enforcement: Present and Future. *IEEE security & Privacy*, 6 n 4, 44-53.
- Price, G. (2003, February 23). Public Key Infrastructure: Challenges and Challengers. *Current development in E-commerce, Lecture Notes*, RHUL.
- Priol, T. (2008). GRID Middleware. *Advanced Grid Research Workshops through European and Asian Co-operation* (pp. 1-11). China: European research consortium for informatics and mathematics.
- Purdue. (2011, February). *NWICG*. Retrieved March 6, 2011, from North West Indiana Grid: <http://www.nwicgrid.org/>
- Rahman, J., & Davis, G. (2009, December 12). *TIME and DIME: Supporting Natural Resource Modelling with .NET and Alchemi CSIRO*. Retrieved March 25, 2012, from <http://alchemi.net/projects.html>
- Rahmeh, O., & Johnson, P. (2008). Towards Scalable and Reliable Grid Networks. *IEEE/ACS International Conference on Computer Systems and Applications (AICCSA-2008)* (pp. 253-259). Doha, Qatar: IEEE Computer Society.
- Ramaswamy, L., Liu, L., & Iyengar, A. (2007, May). Scalable Delivery of Dynamic Content Using a Comprehensive Edge Cache Grid. *IEEE Transactions on Knowledge and Data Engineering*, 614-630.
- Ren, C., Zuo, X., Li, Z., Niu, X., & Yang, Y. (2010). Towards Hierarchical-User RBAC model. *2010 International Conference on Machine Learning and Cybernetics (ICMLC)*, (pp. 2870- 2874). Qingdao, China: IEEE Computer Society.
- Rosado, D. G. (2009, October 24). Development process for Secure Mobile Grid Systems. *PhD Thesis*. Ciudad Real, Spain.
- Roure, D. D., Baker, M. A., Nicholas, R., Jennings, N. R., & Shadbolt, N. (2002). The Evolution of the Grid. *The International Journal computation and Currency : Practice and Experience*.
- Russell, D., & Gangemi, G. (1991). *Computer Security Basics*. Sebastopol, California: O'Reilly Media and Associate.
- Sahai, A., Graupner, S., Machiraju, V., & Moorsel, A. (2003). Specifying and monitoring commercial grids through SLA. *Third IEEE International Symposium on Cluster Computing and the Grid (CCGrid'03)*.
- Samaratim, P., Bertino, E., Ciampichetti, A., & Jajod. (2007). Information Flow Control in Object-Oriented Systems. *IEEE Transactions on Knowledge and Data Engineering*, 9, no. 4, 524-38.
- Sanjeevan, V., Matsunaga, A., Zhu, A., Lam, H., & José, A. B. (2005). A Service-Oriented, Scalable Approach to Grid-Enabling of Legacy Scientific Applications. *IEEE*

- International Conference on Web Services* (pp. 553-560). Orlando, Florida, USA: IEEE Computer Society.
- Sandhu, R., Coyne, E., Feinstein, H., & Youman, C. (1996). Role-based Access Control Models. *IEEE Computer*, 29(2), 38-47.
- Sanga, C., & Venter, M. I. (2010). *A technique for the evaluation of free and open source E-learning systems*. University of the Western Cape, Department of Computer Science. Bellville, Cape Town: University of the Western Cape.
- Shafiq, B., Joshi, J. B., Bertino, E., & Ghafoor, A. (2005, January 27). Secure Interoperation in a Multidomain Environment Employing RBAC Policies. *IEEE Transactions on Knowledge and Data Engineering*, 1557-1577.
- Shen, K., Yang, S., Tian, M., & Liu, P. (2006). Towards a Uniform Monitoring Framework Supporting Interoperability in Grid. *International Conference on Grid and Cloud Computing* (pp. 50-53). IEEE Computer Society.
- Shen, Z.-D., Yan, F., Qiang, W., Wu, X., & Zhang, H. (2006, November 17). Grid System Integrated with Trusted Computing Platform. *IEEE*, 619 - 625.
- Stella, T.-T. (1984). Qualitative research: An overview. *The communication association of the pacific / international communication association convention* (pp. 169-184). Tokyo, Japan: Sage Publications.
- Snelling, F. D., Berghe, S., & Li, Q. V. (2004). Explicit Trust Delegation: Security for Dynamic Grids. *Fujitsu Sci Tech*, 40(2), 282-294.
- Stallings, W. (2002). *Network Security Essentials: Applications and Standards*. Pearson Education, Inc.
- Steve, D. (2009:23, June 23). What is knowledge? Sydney, Australia. Retrieved October 13, 2012
- Stone, A. (2003). *Network Security: Firewalls*. Royal Holloway, University of London: E-commerce infrastructure, Lecture Note.
- Sulistio, A. (2008, May). Advanced Reservation and Revenue-based Resource Management for Grid Systems. *PhD Thesis*. Melbourne, Melbourne, Australia: The University of Melbourne, Australia.
- Tari, Z., & Fry, A. (2001, December). Controlling Aggregation in Distributed Object Systems: A Graph-Based Approach. *IEEE Transactions on Parallel and Distributed Systems*, 23-29.
- Verma, D. C. (2004 ). *Legitimate Applications of Peer-to-Peer Networks*. John Wiley & Sons, Inc.
- Vermeer, M. W., & Apers, P. M. (1996). On the applicability of schema integration techniques to database interoperation. *In International Conference on Conceptual*

- Modeling/the Entity Relationship Approach* (pp. 179-194). Florence, Italy: IEEE Computer.
- Wang, W., Zheng, Y., & Song, G. (2005). The Design and Implementation of Scalable Information Services in a Grid Environment. *IEEE International Conference on Services Computing* (pp. 265-267). Orlando, Florida, USA: IEEE Computer Society.
- Wang, S., Shook, E., Padmanabhan, A., Briggs, R., & Pearlman, L. (2006). Developing the Modular Information Provider (MIP) to Support Interoperable Grid Information Services. *International Conference on Grid and Cloud Computing* (pp. 448-453). IEEE Computer Society.
- Watkins, R., Mark, M., Leonard, T., & Surridge, M. (2007). Cross-middleware Interoperability in Distributed Concurrent Engineering. *International Conference on e-Science and Grid Computing* (pp. 561-568). Bangalore, India: IEEE Computer Society.
- Weihua, L., & Shixian, L. (2005). Using Information-Flow Theory to Support Information Interoperability in the Grid. *International Conference on Information Technology and Applications* (pp. 272-275). Sydney, Australia: IEEE Computer Society.
- Welch, D., & Lathrop, S. (2003). Wireless Security Threat Taxonomy. *Proceedings of the 2003 IEEE workshop on information assurance United States Military Academy West Point, NY*.
- ZeFeng, C., HaiCheng, Y., & Rong, M. (2010). Interoperability of Digitizing Systems based on Semantic Grid. *2010 International Conference on E-Business and E-Government* (pp. 1434-1436). IEEE Computer Society.
- Zhang, G., & Parashar, M. (2006). SESAME: Scalable, Environment Sensitive Access Management Engine. *Journal Cluster Computing*, 19 - 27.
- Zhang, X., Wu, H., & Wu, Z. (2007). Scalable Collaborative Virtual Environment based on Open Grid Services Architecture. *International Conference on Semantics, Knowledge and Grid* (pp. 580-581). Beijing, China: IEEE Computer Society.
- Zhao, G., & Chadwick, D. W. (2008). On the Modeling of Bell-LaPadula Security Policies using RBAC. *Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE '08* (pp. 257-262). Rome: IEEE.
- Zhou, J., Deng, R. H., & Bao, F. (1999). Evolution of Fair Non-repudiation with TTP. *Proceedings of the 4th Australasian Conference on Information Security and Privacy, UK ACISP '99*. Springer-Verlag London.

## Appendices

### APPENDIX A:

#### *Findings of the content analysis*

#### **Access Control List (ACL)**

##### *Implementation*

ACL remains the earliest, oldest as well as the most basic form of access control (Stone, 2003). It became prominent in the early 70s to constrain access to data and files on distributed systems (Roure, Baker, Nicholas, Jennings, & Shadbolt, 2002) due to the emergence of multi user operating systems. The functioning of ACL is simple: each of the resources on the system has a corresponding list of entities that may access the resource and the associated actions that such an entity can perform on the resource.

For example, a data structure can hold a list of authorized users that may access a particular file on the file system. A flag (Levine, 2002.) specifies whether a user can read, execute, modify, delete or write a file or perform some combination of two or more operations on the file (Foster I. , Kesselman, Tsudik, & Tuecke, 1998).

Thus an ACL, is a list of genuine and authorized users as well as their rights and privileges for each of the resources in a virtual organization (Houmb., Georg, France, & Matheson, 2004).

The access control that governs the access to resources in domain A (see Figure 42) can be expressed as follows:

$$ACL (RA) = [(UC, [opA1, opA2, opA3.....]), (UD, [opA5, opA6, opA7.....]).....]$$

Where opA1.....opA7 are operations that are allowed in domain A. RA are the resources in domain and UC and UD are the number of users in domain C and D respectively.

### *Strengths of ACL*

Three main strengths of the ACL were identified, (using content analysis). They include the following:

- It provides a direct and straightforward means of denying or granting access for specific group of users. This is achieved by verifying the Access Control List to know whether a group is in the list or not.

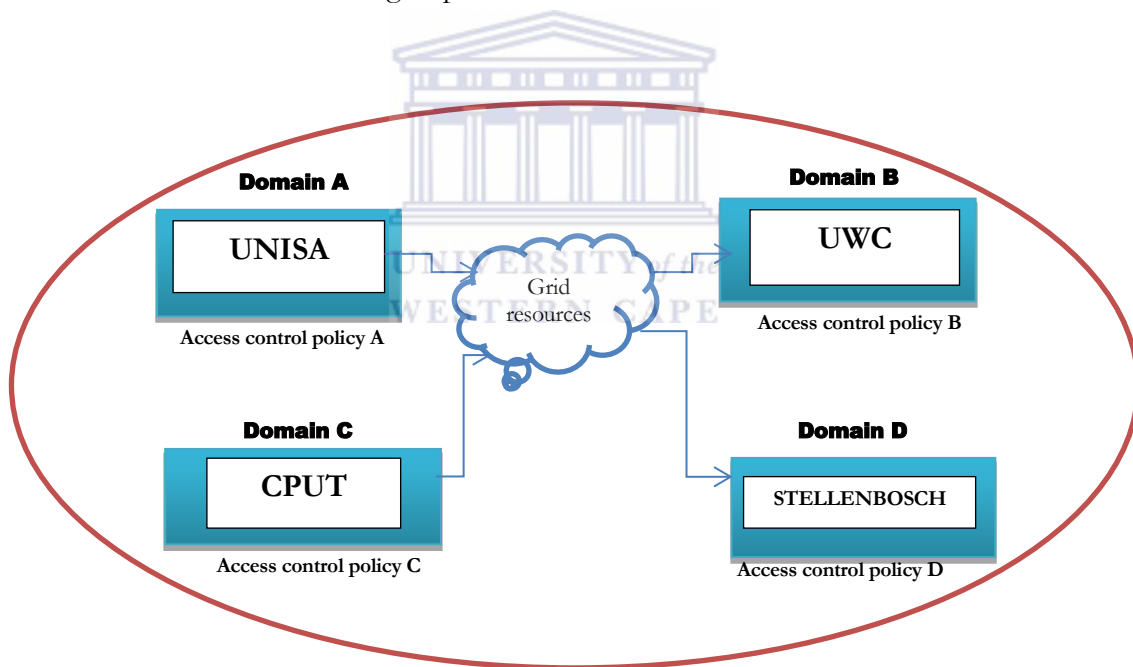


Figure 42: *Multiple-domain resource sharing environment*

- ACL allows simplicity and avoids complexity in granting access to the user to resources (Hu, Ferraiolo, & Kuhn, 2006)
- Some commercial operating systems such as Microsoft & Windows XP and Windows 2007 employ ACL for their built-in access control

### *Weaknesses of ACL*

The weakness of ACL include lack of interoperability, it is static and vulnerability:

- Lack of interoperability. The various domains and organisations which are involved in a virtual organisation (VO) need to agree on what permission and access framework to be implemented. This is to be done at an advance stage so that each domain administrator can easily enforce the same policy. However, it might be difficult for organisations to agree on all policies.
- ACL is static. The ACL is strictly tied to resources and does not allow for scalability. The individual's user account is managed by the domain administrator and is managed for each resource. Whenever a user joins or leaves the grid, the domain administrator in each of the VO's will have to update the ACL of the resource by deleting the access of the user or adding rights to the resource for the user.
- ACL is very vulnerable to errors. Since the grid is well known for its dynamic nature, the implementation of an ACL on each individual domain poses a great challenge. At any time when a user is withdrawn from a project, the domain administrator will have to go through each of the resources to remove all the access privileges granted to such a user. When revoking access rights, some of the privileges may be overlooked.

### **Mandatory access control (MAC)**

This type of access control policy allows decisions to be made over and above the directive and control of each of the resource's owner. There is a central authority that decides and determines which resources and information should be accessed by whom (agent) and the access rights of the user cannot be manipulated or altered (Imine, Cherif, & Rusinowitch, 2009). This security model allows every user and object to be assigned a label called a sensitivity label, which comprises compartments and a level of secrecy (Azeez, Iyamu, & Venter, 2011).

Four tiers of security levels are considered in the following order: TOPSECRET > SECRET > CONFIDENTIAL > UNCLASSIFIED (Haidar, 2002).

SECRET is considered to be the sensitivity label for the logistics file while SECRET implies the security level. Also, the sensitivity label of a file for *user* is *TOPSECRET* where *TOPSECRET* indicates the security level.

Mandatory access control is based on two main principles as formulated and stated by (Pedayachee, 2009). They are:

- There is no read-up: This implies that a grid user running an application at “Secret” level would not be permitted to access (read, write, delete) any file labelled “Top Secret”. This is because Top Secret > Secret in the security hierarchy. A subject is only allowed to read a dominated object in an access class.
- There is no write-down: In this case, whenever a grid user runs a process labelled “Secret” such a user should not be allowed to access (write) any file tagged “Confidential”. A subject can only write to objects that have access dominating that of the subject. This rule is called the “\*-property”

#### *Strengths of MAC*

Four main strengths of the MAC were identified:

- The decision about the access control policy is taken by the central authority hence; each grid user cannot alter or change the access rights and privileges on his own
- It adopts a labelling or tagging mechanism with a set of interfaces to enforce protection decision. For example, a grid user running a process at the “*secret*” level will be disallowed from accessing file tagged “*Top Secret*”. This is termed as “simple security rule”
- It provides a greater and reliable level of protection to the resources when compared to access control list model



- It does not subscribe to illegal flow of information which occurs when a message is transmitted from one end to the other, contradictory to information flow security policy (Samaratim, Bertino, Ciampichetti, & Jajod, 2007).

Mandatory access control is known for its numerous advantages (Pfleeger, 1997) especially when compared with discretionary access control, however the model still has many deficiencies (Russell & Gangemi, 1991).

#### *Weaknesses of MAC*

- Information downgrading: this occurs when information is transferred from a higher security level to a lower level
- Adopting a MAC model may lead to a large chunk of the operating system being incorporated into the Trusted Computing Base (TCB)
- MAC creates a large semantic gap between the access control of programming languages and the operating systems. Hence languages like java virtual machine (JVM) cannot implement mandatory access control (Pedayachee, 2009).

The adoption and implementation of MAC models were prevalent in military applications, such as role specification for the military while in a battle field. According to Anderson (2001), the model is presently being adopted in systems that require intricate relationships among the various entities, for example in medical data applications which involve sensitive data as well as e-business applications which require stringent privacy control mechanisms (Samaratim, Bertino, Ciampichetti, & Jajod, 2007).

#### **Distributed Authorization Model (DAM)**

This is considered to be a user-centered authorization oriented model (Kealey & Protheroe, 1996). The motive behind this security model is that each of the domains participating in the virtual organization project has a representative or proxy, tasked with the responsibility of translating a global credential to a local one. Therefore, each of the grid users is expected

to possess his/her account from his immediate administrative domain (McCroskey & Richmond, 1990). However, the access right of each of the users is managed and decided on by the virtual organization domains. Before any access can be granted to a resource on the grid, a mapping of the credentials of the user's administrative domain must be compared with that of the resource domain.

Consider Figure 1, where the decision of the access privileges of user in domain C (UC) to access resource of domain D (R.D) is ascertained and achieved through mapping of UC's identity from domain C (i.e Public Key Infrastructure (PKI) (Price, 2003) to a local account in administrative domain D (i.e Operating System 2 (OS2)). As a result of the mapping of these two different domains, U.C can easily access RD as a local and immediate user in domain D. Also, domain D can easily apply its domain security policy. This security solution has been adopted by various popular grid projects like Unicore and Globus (Haidar, 2002).

#### *Strengths of DAM*

- The local site director and administrator are in full control of their resources.
- It is transparent since both the passwords and usernames are provided with the subject's name clearly written on the certificate. Hence it is very easy to find someone who carries out an unauthorized action on a particular resource.
- Changes in the security framework are not required in the local domain. The security protocol can be established with role base access control (RBAC) mechanisms.

#### *Weaknesses of DAM*

- Each of the domains on the grid depend directly on the central operation;
- User's access rights and privileges are controlled , managed and monitored on each virtual organisation (VO) sites (Haidar, 2002);
- Problem of scalability: grid user should have local accounts in all the VO sites where he should access resources. Hence, the administrator will have to update a user's

account and schedule its access right whenever a new grid user is added or when he has his duty reviewed;

- Problem of inconsistency: deleting user's access right is complex and cumbersome.

All the system administrators must be well informed should in case a grid user change or when there is a change in policy. If anyone of the VO administrators is not acquainted, there will be inconsistency in the policy.

#### *The functioning of DAM explained*

Consider Figure 43 and Figure 44 where two domains (with different security policies and protocols) are involved; J and K. When a user from either domain requests to access a resource from the other domain, the user will have to follow the security requirements of that domain before access will be granted. If the user is unable to do so, access will be denied. Figure 43 shows identity mapping using ACL. Anytime a user in domain J wants to access resource in domain K (R.K), the following situations hold:

- User in domain J (UJ) sends across through stage 1, his X509 (ITU-T approved standard for specifying what information goes into a certificate and clearly explains its written format) certificate to resource in domain K. This contains his subject name as in stage 2. The resource confirms the validity of the certificate to know whether there is a private key with the sender which tallies to the public key of the certificate
- If the above (authentication) succeeds, the subject name is extracted (SN. UJ) from the certificate as shown in 3.
- The resource is thereafter compares the subject name to various entries already available in the mapping database at phase 4. The feedback could either be True (authorised) or false (not authorised)
- If the feedback is false (not authorised), it implies that there is no match and that the user is not allowed to access the resource

- But if the feedback is True (authorised), both the password and the user name that are corresponding to subject name from the database are fetched by the resource as appears in stage 5. At this stage, the user can become accessible to resource as a local (grid user) in domain K..

In Figure 44 the process of identity mapping using RBAC in a distributed authorisation is shown. All conditions mentioned under identity mapping using ACL Figure 42 hold, except that in stage two, the resource has to extract the subject name (SN, UJ) from the certificate based on the role specified for the user

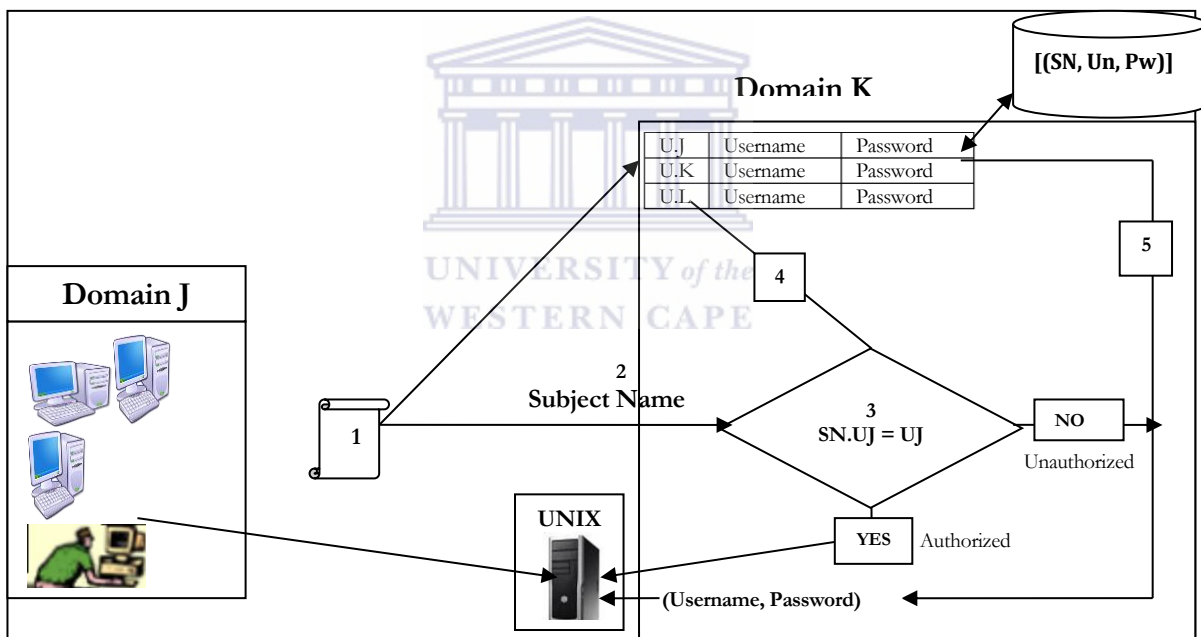


Figure 43: Process of identity mapping using Access Control List

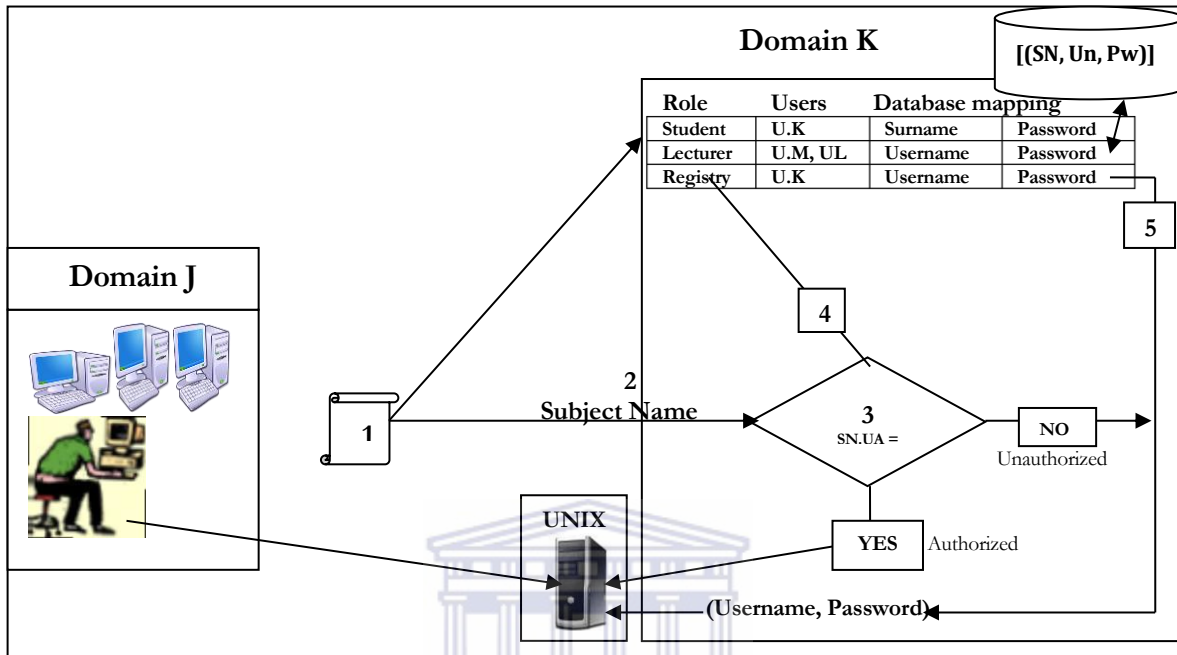


Figure 44: Process of identity mapping using role based access control

Once this is achieved and established, the grid user can access resources as a local user in domain K as if he or she is a local user.

### Context Based Security Model (CBSM)

This is a new security model designed to handle security challenge emanating from the high mobility of pervasive computing and different devices used in these types of environment (Etalle & Winsborough, 2007). This model handles context as the most important and basic principle in policy enforcement and specification. The context based security model presents the contexts of the various security policies and how they are associated as well as which agents may operate within these policies. It specifies the action which must be carried out and those that are forbidden on resources in contexts. It associates different agents with the context dynamically (Ferreira, et al., 2006).

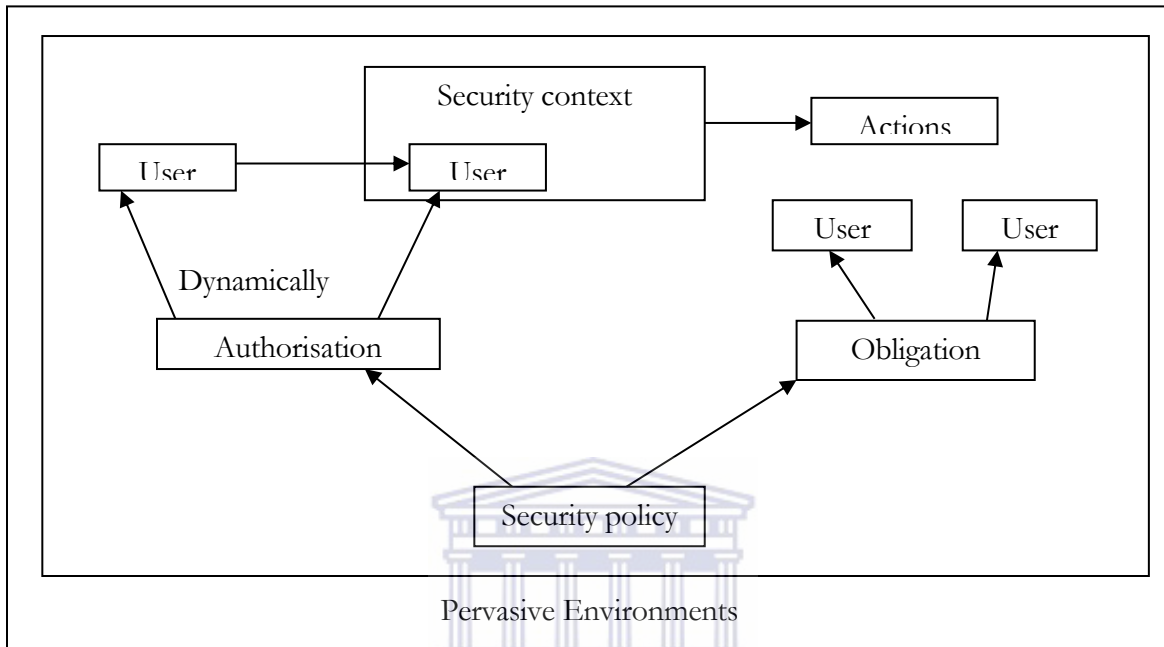


Figure 45: Context-based Security Policy (Pretschner, Hilty, Schutz, Schaefer, & Walter, 2008).

This security model makes use of contextual graphs to specify various security policies. It is used as a tool which eases security specification for complex domains with various heterogeneous devices and services (Georgiev & Georgiev, 2001). The application and exploitation of context for merging security policies eases and simplifies access control management policies by enhancing the reuse of policy specification and by simplifying complexities of policy update and revocation (Frey, Tannenbaum, Foster, & Livny, 2001). This security policy is unique in the sense that it allows policy specifications and enforcement as well as fine-grained access control (D'Arcy & Hovav, 2007).

*Security context* can be defined as the information gathered from the application and the user's environments considered to be very important and useful to the security of both the application and user (Gouglidis & Mavridis, 2009). To build a security context, some

contextual information may be used, such as user's location, his preference, user's identity, information sensitivity etc.

*Security policy* on the other hand is a specification that states clearly and precisely the authorized practices with all the denials for each user in each context.

Figure 42 shows a conceptual model of a context based security policy defined as a union of actions in a grid based environment and context. The pervasive grid environment is firstly controlled and monitored with a particular configuration of the security policy in initial context (Patrick & Ghita, 2004). This context is dynamically changing in request to variation in the environment. The security policy should adjust itself to any new context so as to bridge and fill the breaches that might be available by the new context ( Baktash, Karimi, Meybodi, & Bouyer, 2010). When a context is finally established in the environment, then appropriate action follows that is when an authorized user approaches a particular security context, such context will automatically be associated with the corresponding action. Context can be any relevant information like the user's location, the features of the underlying device and many others (Dewan, Grundin, & Horvitz, 2007). Despite numerous benefits of this security model, it is new and requires more testing to confirm its reliability in a grid based environment.

#### *Strengths of CBSM*

- It caters for adequate and efficient security challenges in a pervasive grid based applications and environment
- It offers a security solution based on the explicit specification on the context (situation) of use of the system
- It conveniently implements security policy in an environment where its policies are considered static

*Weaknesses of CBSM*

- CBSM is new and its application has not been adequately explored hence its efficiency and reliability cannot be fully guaranteed
- It is neither flexible nor interoperable since its application is limited to finding a security solution to a pervasive computing environment.
- It has limited areas of application

**Role based access control (RBAC)**

The major problem with some of the previously explained security models is administering individuals with roles. To solve this anomaly, RBAC authorization model could be adopted. Role based access control came into existence in the early 90s as a form of security model for enforcing and managing security policies in a virtual organisation (Herath & Rao, 2009).

RBAC has a focus on the users and the jobs they perform within a given organisation. Hence, it is useful and appropriate for a dynamic and large-scale environment such as the grid. The main opinion of RBAC is that roles are associated with permissions (Imine, Cherif, & Rusinowitch, 2009) and users are assigned appropriate roles or duties. With RBAC model, only genuine and authorized users are permitted to access various resources on the grid. This provides simplicity for authorization management when providing chance for dynamism in enforcing and specifying certain enterprise security policies (Mao, Li, Chen, & Jiang, 2009).

Figure 46 reveals the relationships among different users (A, B, C, D,E, F, etc.), different roles (R1, R2, R3, R4, R5, R6,etc.) and permissions (P1, P2, P3, P4, P5, P6, etc.) in a role based access control model. Each user can be assigned to a role (or several roles) and a role can be assigned to many users as shown in the Figure 46. In essence, there are one-to-one, many-to-one and one-to-many relationships between role and user. Each user has many or one permission and likewise permissions could be associated and assigned to just one or many roles according to the security policy. Finally, T1, T2, T3, T4, T5, T6, etc., represent



database of resources or data, and its access to the grid. Permission is either area right or a write right on the grid.

*Strengths of the RBAC model*

- Reduction in both administrative complexity and cost.

With role based access control model, the local domain administrator can provide permissions to roles (rather than individuals), according to local and global policies ( Chadwick & Otenko, 2002).

$$RBAC (UA) = [Dom1, Role1), (Dom2, Role2).....] \dots\dots\dots equation 1$$

The above implies that applying RBAC for user A specifies or defines Role1 for Domain 1 and Role 2 for Domain 2.

$$Permission (Dom1, Role1) = [Opr11 (Writing), Opr12 (deleting), Opr13 (updating).....] \dots\dots\dots equation 2$$

The above means specified permissions for Domain 1 and Role1 are given as operation11 (Writing), operation12 (deleting) and operation13 (updating) to be carried out.

- Operational consistency

Each role can be associated with a position and as such, a set of operations can be allowed. For example, three lecturers with similar roles will have the same permissions (Haidar, 2002).

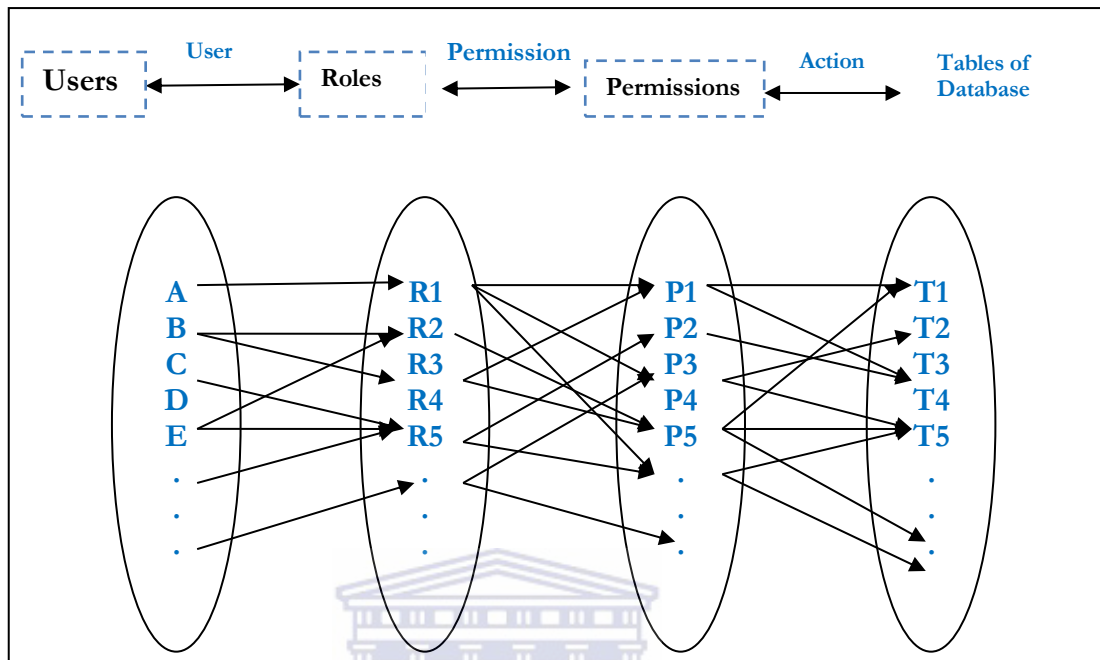


Figure 46: User-Role-Permission mapping in a flexible RBAC

- RBAC can incorporate other models such as MAC and DAC ( Pedayachee, 2009).
- RBAC can be used with various applications that are interoperable, flexible and scalable.

#### *Weaknesses of the RBAC model*

- With RBAC, allocation of files and servers may not be compatible with any establishment structure that needs and requires users to focus on empirical matters like payment of bills and opening of account
- It is too difficult and costly to implement and achieve the least privilege condition since it is complex to tailor access based on different constraints and attributes
- User access to resources are only regulated mainly on the basis of role (the activities the users perform) and thus derecognize some other important indices for accessing resources

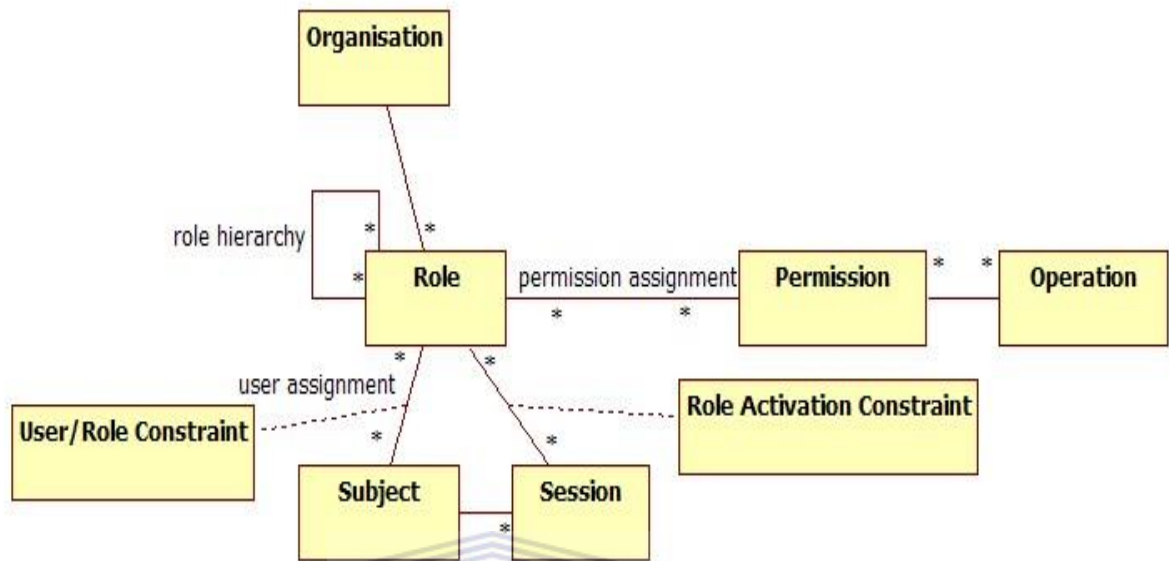


Figure 47: Relationship of elements in a flexible RBAC  
 (Source: <http://en.wikipedia.org/wiki/File:rbac.jpg>)

The formal specification, definition and explanation of individual elements in a flexible RBAC are shown in Figure 47. The abbreviations are explained as follows: The star (\*) sign in Figure 47 implies many-to-many relationship between the entities.

Table 9: Description of elements in a flexible rbac

Abbreviation	Name	Description
S	Subject	A person or agent involved
R	Role	Job definition or specification
P	Permissions	It specifies a mode of accessing a resource
SE	Session	A mapping of Subject (S), Role (R), and / or Permission (P)
SA	Subject assignment	Assigns a subject to a role
RH	Role Hierarchy	The position / level of a role
PA	Permission Assignment	Gives permission to a role

In set theory notation the relationships are shown as follows:

$PA \subseteq P \times R$ : This is a many to many permission to role assignment relation.

$SA \subseteq S \times R$  This is a many to many subject to role assignment relation.

$RH \subseteq R \times R$  : This is a partial order on R referred to as role dominance relation or role hierarchy.

### RBAC Framework model

RBAC framework model divides RBAC into four main categories: core RBAC, hierarchical RBAC, static separation of duty relation and dynamic separation of duty relations. The function of each of the categories is presented in Table 10

Table 10: RBAC Framework model

Category	Function
Core RBAC	<ul style="list-style-type: none"> <li>Introduces the concept of role activation as part of a user's session within a computer system.</li> <li>Required in any RBAC system, but the other components are independent of each other and may be implemented separately.</li> </ul>
Hierarchical RBAC	<ul style="list-style-type: none"> <li>Relations for supporting role hierarchies (inheritance among roles)</li> </ul>
Static Separation of Duty Relations	<ul style="list-style-type: none"> <li>Adds exclusivity relations among roles w.r.t. user assignments</li> <li>Potential for inconsistencies w.r.t. static separation of duty relations and inheritance relations of a role hierarchy</li> <li>Defines relations in both the presence and absence of role hierarchies.</li> </ul>
Dynamic Separation of Duty Relations	<ul style="list-style-type: none"> <li>Exclusivity relations w.r.t. roles that are activated as part of a user's session</li> </ul>

APPENDIX B

Tables containing information for databases of university, hospital and banking

Table 11: University database for domain A

ID No	Student NO	Surname	First name	Initial	DOB	Passport No
1000	3008814	Azeez	Nureni	N.A	1978/07/10	QOO12345
1001	3066278	Adewale	Abiola	M.D	1980/04/08	RE0047476
1200	2340078	Abidoye	Philip	P.O	1976/06/09	WE345678
1220	2357858	Scholtz	Josue	S.J	1981/04/05	VB7878784
1260	3400993	Magnuth	Henry	H.M	1975/06/09	FD8785758
1320	3476002	Andy	Liu	X.L	1984/02/09	RE7878784
1400	3455266	Achmed	Imran	I.A	1986/02/09	UD5785785
1523	2004556	Jonathan	Magnus	I.M	1974/07/10	TY8989956

Table 12: University database for domain A (Continued from Table 10)

National	Residential	Gender	MArital stat	Home langu	Postal Address
Nigerian	Belhar	Male	Married	Yoruba	Dept of Computer Science
Nigerian	HPR	Male	Married	Yoruba	Dept of Computer Science
China	HPR	Male	Single	Chineese	Dept of Computer Science
South Afric	Cassinga	Female	Single	Africaans	Dept of Computer Science
Congo RC	FBR	Male	Single	Samdaba	Department of Physics
Namibia	DOS	Female	Single	Shaulli	Department of Maths
Ghana	Cassinga	Male	Single	Huoity	Department of Regious Studies
Cameroun	SANTOS	Female	Married	Gonghil	Department of Geology

Table 13: University database for domain A (Continued from table 11)

Tel no	Email	Disability	Next of kin	Examination Auth	School	Last school
0738991735	nurayhn@yahoo.ca	None	Abdul Waahid	WAEC	UWC, RSA	Ede
0713456878	abidoeye.ade@gmail.com	None	Ademola	SSCE	UWC, RSA	Ilesha
0756355252	andy@gmial.com	None	Xium	MATRIC2003	UWC, RSA	Njoku Lee
0723453545	imran@gmail.com	None	Nabila	GRADE 10	UWC, RSA	Cape Town
021(0)56156	hjuj@gmail	Hearing	Joy	GRADE 10	UWC, RSA	Poskia
0749988452	wale@gmail.com	None	Shola	GRADE 2	UWC, RSA	Somekia
021(0)67476	schortz@gmail.com	Astigmatism	Jerry	GRADE10	UWC, RSA	Hausa
0746767790	henry@gmail.com	None	Mercy	GRADE 4	UWC, RSA	Igbo

Table 14: University database for domain A (Continued from Table 12)

Passport expiry dat	permit num	permit type	Medical aid	Validity of med	postal code	course	department
2014	23KIO456	Study	INGWE	March, 2013	7535	PhD	Computer Science
2014	45UYO045	Study	MOMENTUM	April, 2013	7556	PhD	Computer Science
2013	66UYO041	Study	NETCARE	2013	7534	PhD	Biochemistry
Not Applicable	Not Applicable	Not Applicable	MEDIMED	Not Applicable	6573	Honours	Computer Science
2015	67KIO7834	Study	INGWE	March, 2015	5534	Masters	Philosophy
2013	76HG6879	Study	INGWE	Dec., 2012	7887	Masters	Mathematics
2015	67DA9975	Study	MAKOTI	Nov., 2012	7535	Masters	Religious studies
2013	87FG12235	Study	MALCOR	2014	6778	PhD	Geology

Table 15: Hospital database for domain B

Patient_ID	File No	Surname	First name	Patient_Condition	Date_Of_Adm
1000	0031	Azeez	Nureni	Medical Checkup	2012/09/04
1200	0045	Abidoeye	Philip	Eye problem	2012/06/04
1320	0067	Andy	Liu	Hearing problem	2012/03/06
1400	0012	Achmed	Imran	Back Pain and X-ray	2011/09/04
1523	0056	Jonathan	Magnus	Pregnancy	2012/01/02
1001	0023	Adewale	Abiola	Blood Test	2012/09/09
1220	0013	Scholtz	Josue	Car accident	2011/09/01
1260	0011	Magnuth	Henry	Head injury	2012/03/06
*					

Table 16: Hospital database for domain B (Continued from Table 14)

Insurance_Num	Province	Admission Unit	Employment status	Patient Account
009	The Western Cape	Out-Patient	Full-time Position	001WF
007	Eastern Cape	Optometry	Full-time Position	985TY
003	Northern Cape	Out-Patient	Part-time Position	686YU
002	Limpopo	Radiology	Full-time Position	345TR
008	Cape Town	Ante natal	Full-time Position	788HG
005	The Free State	Medical Lab	Part-time Position	682FD
004	KwaZulu-Natal	Emergency	Part-time Position	242DS
006	Mpumalanga	Emergency	Part-time Position	909ED

Table 17: Hospital database for domain B (Continued from Table 15)

Health provider	Age	Bloodgroup	Genotype	Race
INGWE	23	A	AA	BLACK
MOMENTUM	34	B	AS	COLOURED
NETCARE	27	O	AA	INDIA
MEDIMED	29	AB	SS	BLACK
INGWE	20	AB	AC	WHITE
INGWE	28	O	SC	INDIA
MAKOTI	25	B	AA	BLACK
MALCOR	32	A	AC	WHITE

Table 18: Banking database for domain C

ID_NO	Customer Name	service code	Account ID	Current balance
1000	Azeez, N.A	FD9989	45454599XX	R 56.5XX
1200	Abidoeye, P.O	YU7878	57757577XX	R 100XX
1320	Andy, X.L	HJ7880	47747744XX	R 562XX
1400	Achmed, I.A	WE4545	67677676XX	R 00.7XX
1523	Jonathan, I.M	QW576	56565655XX	R 06.7XX
1001	Adewale, M.D	NH7676	86868612XX	R 129XX
1220	Scholtz, S.J	YE85852	67676768XX	R 451XX
1260	Magnuth, H.M	BG2323	13243535XX	R 000XX

Table 19: Banking database for domain C (Continued from Table 17)

Initial Date of Transacti	Type of Account	Tax_ID
Jan 16, 2012	Saving	0011
March 10, 2012	Cheque	0056
Feb 19, 2012	Saving	0085
Dec 18, 2011	Saving	0045
Sept 15, 2012	Cheque	0034
June 12, 2012	Cheque	0023
April 10, 2012	Cheque	0057
Aug 10., 2012	Saving	0078

APPENDIX C

Report of cross domain query of Figure 41 for domains A and B

Domain A:and B ( University database and Hospital database)

ID No	Student NO	Surname	First name	Initial	DOB	Passport No	Nationality	Residential Area	Gender	MArital status	Home language	Postal Address	Tel no	Email	
1000	3008814	Azeez	Nureni	N.A	1978/07/10	QOO12345	Nigerian	Belhar	Male	Married	Yoruba	Dept of Computer Science	0738991735	nurayhn@yahoo.ca	
1001	3066278	Adewale	Abiola	M.D	1980/04/08	RE0047476	Namibia	DOS	Female	Single	Shaulli	Department of Maths	0749988452	wale@gmail.com	
1200	2340078	Abidoye	Philip	P.O	1976/06/09	WE345678	Nigerian	HPR	Male	Married	Yoruba	Dept of Computer Science	0713456878	abidoye.ade@gmail.com	
1220	2357858	Scholtz	Josue	S.J	1981/04/05	VB7878784	Ghana	Cassinga	Male	Single	Huoity	Department of Regious Studies	021(0)67476	schortz@gmail.com	
1260	3400993	Magnuth	Henry	H.M	1975/06/09	FD8785758	Cameroun	SANTOS	Female	Married	Gonghil	Department of Geology	0746767790	henry@gmail.com	
1320	3476002	Andy	Liu	X.L	1984/02/09	RE7878784	China	HPR	Male	Single	Chinese	Dept of Computer Science	0756355252	andy@gmial.com	
1400	3455266	Achmed	Imran	I.A	1986/02/09	UD5785785	South Africa	Cassinga	Female	Single	Africaans	Dept of Computer Science	0723453545	imran@gmail.com	
1523	2004556	Jonathan	Magnus	I.M	1974/07/10	TY8989956	Congo RC	FBR	Male	Single	Samdaba	Department of Physics	021(0)56156	hjuy@gmail	
Disability	Next of kin	Examination Authority	School	Last school year status	Passport expiry date	permit number	permit type	Medical aid name	Validity of medical aid	postal code	course	department			
None	Abdul Waahid	WAEC	UWC, RSA	Ede	2014	23KIO456	Study	INGWE	March, 2013	7535	PhD	Computer Science			
None	Shola	GRADE 2	UWC, RSA	Somekia	2013	76HG6879	Study	INGWE	Dec., 2012	7887	Masters	Mathematics			
None	Ademola	SSCE	UWC, RSA	Ilesha	2014	45UYO045	Study	MOMENTUM	April, 2013	7556	PhD	Computer Science			
Astigmatism	Jerry	GRADE10	UWC, RSA	Hausa	2015	67DA9975	Study	MAKOTI	Nov., 2012	7535	Masters	Religious studies			
None	Mercy	GRADE 4	UWC, RSA	Igbo	2013	87FG12235	Study	MALCOR	2014	6778	PhD	Geology			
None	Xium	MATRIC2003	UWC, RSA	Njoku Lee	2013	66UYO041	Study	NETCARE	2013	7534	PhD	Biochemistry			
None	Nabila	GRADE 10	UWC, RSA	Cape Town	Not Applicable	Not Applicable	Not Applicable	MEDIMED	Not Applicable	6573	Honours	Computer Science			
Hearing	Joy	GRADE 10	UWC, RSA	Poskia	2015	67KIO7834	Study	INGWE	March, 2015	5534	Masters	Philosophy			
Patient_ID	File No	Surname	First name	Patient_Condition	Date_Of_Admission	Insurance_Num	Province	Admission Unit	Employment status	Patient Account	Health provider	Age	BloodGroup	Genotype	Race
1000	0031	Azeez	Nureni	Medical Checkup	2012/09/04	009	The Western Cape	Out-Patient	Full-time Position	001WF	INGWE	23	A	AA	BLACK



## APPENDIX C

*Report of cross domain query of Figure 41 for domains A and B*

Domain A:and B ( University database and Hospital database)

ID No	Student NO	Surname	First name	Initial	DOB	Passport No	Nationality	Residential Area	Gender	MArital status	Home language	Postal Address			Tel no		Email
1001	0023	Adewale	Abiola		Blood Test	2012/09/09	005	The Free State		Medical Lab	Part-time Position	682FD	INGWE	28	O	SC	INDIA
1200	0045	Abidoeye	Philip		Eye problem	2012/06/04	007	Eastern Cape		Optometry	Full-time Position	985TY	MOMENTUM	34	B	AS	COLOURED
1220	0013	Scholtz	Josue		Car accident	2011/09/01	004	KwaZulu-Natal		Emergency	Part-time Position	242DS	MAKOTI	25	B	AA	BLACK
1260	0011	Magnuth	Henry		Head injury	2012/03/06	006	Mpumalanga		Emergency	Part-time Position	909ED	MALCOR	32	A	AC	WHITE
1320	0067	Andy	Liu		Hearing problem	2012/03/06	003	Northern Cape		Out-Patient	Part-time Position	686YU	NETCARE	27	O	AA	INDIA
1400	0012	Achmed	Imran		Back Pain and X-ray	2011/09/04	002	Limpopo		Radiology	Full-time Position	345TR	MEDIMED	29	AB	SS	BLACK
1523	0056	Jonathan	Magnus		Pregnancy	2012/01/02	008	Cape Town		Ante natal	Full-time Position	788HG	INGWE	20	AB	AC	WHITE

*Cross-domain queries reports*

*B. Report of cross domain query of Figure 42 for domains A and C*

**Domain A and C:(Hospital Database and Banking databases)**

Patient_ID	File No	Surname	First name	Patient_Condition	Date_Of_Admission	Insurance_Num	Province	Admission Unit	Employment status	Patient Account	Health provider	Age	BloodGroup	Genotype	Race
1000	0031	Azeez	Nureni	Medical Checkup	2012/09/04	009	The Western Cape	Out-Patient	Full-time Position	001WF	INGWE	23	A	AA	BLACK
1001	0023	Adevale	Abiola	Blood Test	2012/09/09	005	The Free State	Medical Lab	Part-time Position	682FD	INGWE	28	O	SC	INDIA
1200	0045	Abidoeye	Philip	Eye problem	2012/06/04	007	Eastern Cape	Optometry	Full-time Position	985TY	MOMENTUM	34	B	AS	COLOURED
1220	0013	Scholtz	Josue	Car accident	2011/09/01	004	KwaZulu-Natal	Emergency	Part-time Position	242DS	MAKOTI	25	B	AA	BLACK
1260	0011	Magnuth	Henry	Head injury	2012/03/06	006	Mpumalanga	Emergency	Part-time Position	909ED	MALCOR	32	A	AC	WHITE
1320	0067	Andy	Liu	Hearing problem	2012/03/06	003	Northern Cape	Out-Patient	Part-time Position	686YU	NETCARE	27	O	AA	INDIA
1400	0012	Achmed	Imran	Back Pain and X-ray	2011/09/04	002	Limpopo	Radiology	Full-time Position	345TR	MEDIMED	29	AB	SS	BLACK
1523	0056	Jonathan	Magnus	Pregnancy	2012/01/02	008	Cape Town	Ante natal	Full-time Position	788HG	INGWE	20	AB	AC	WHITE

ID_NO	Customer Name	service code	Account ID	Current balance	Initial Date of Transaction	Type of Account	Tax ID
1000	Azeez, N.A	FD9989	45454599XX	R 56.5XX	Jan 16, 2012	Saving	0011
1001	Adevale, M.D	NH7676	86868612XX	R 129XX	June 12, 2012	Cheque	0023
1200	Abidoeye, P.O	YU7878	57757577XX	R 100XX	March 10, 2012	Cheque	0056
1220	Scholtz, S.J	YE85852	67676768XX	R 451XX	April 10, 2012	Cheque	0057
1260	Magnuth, H.M	BG2323	13243535XX	R 000XX	Aug 10, 2012	Saving	0078
1320	Andy, X.L	HJ7880	47747744XX	R 562XX	Feb 19, 2012	Saving	0085
1400	Achmed, I.A	WE4545	67676766XX	R 00.7XX	Dec 18, 2011	Saving	0045
1523	Jonathan, I.M	QW576	56565655XX	R 06.7XX	Sept 15, 2012	Cheque	0034

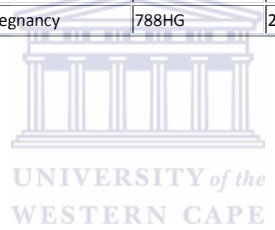
Domain A: University Database		
ID No	Surname	Nationality
1000	Azeez	Nigerian
1001	Adewale	Namibia
1200	Abidoeye	Nigerian
1220	Scholtz	Ghana
1260	Magnuth	Cameroun
1320	Andy	China
1400	Achmed	South Africa
1523	Jonathan	Congo RC

Domain B: Hospital Database			
File_No	Patient_Condition	Patient_Account	Age
0031	Medical Checkup	001WF	23
0023	Blood Test	682FD	28
0045	Eye problem	985TY	34
0013	Car accident	242DS	25
0011	Head injury	909ED	32
0067	Hearing problem	686YU	27
0012	Back Pain and X-ray	345TR	29
0056	Pregnancy	788HG	20

Domain C: Banking Database		
service code	Account_ID	Tax_ID
FD9989	45454599XX	0011
NH7676	86868612XX	0023
YU7878	57757577XX	0056
YE85852	67676768XX	0057
BG2323	13243535XX	0078
HJ7880	47747744XX	0085
WE4545	67677676XX	0045
QW576	56565655XX	0034

*C. Report of cross domain query of Figure 43 for domains A, B and C*

ID No	Surname	Nationality	File_No	Patient_Condition	Patient_Account	Age	service code	Account_ID	Tax_ID
1000	Azeez	Nigerian	0031	Medical Checkup	001WF	23	FD9989	45454599XX	0011
1001	Adewale	Namibia	0023	Blood Test	682FD	28	NH7676	86868612XX	0023
1200	Abidoeye	Nigerian	0045	Eye problem	985TY	34	YU7878	57757577XX	0056
1220	Scholtz	Ghana	0013	Car accident	242DS	25	YE85852	67676768XX	0057
1260	Magnuth	Cameroun	0011	Head injury	909ED	32	BG2323	13243535XX	0078
1320	Andy	China	0067	Hearing problem	686YU	27	HJ7880	47747744XX	0085
1400	Achmed	South Africa	0012	Back Pain and X-ray	345TR	29	WE4545	67677676XX	0045
1523	Jonathan	Congo RC	0056	Pregnancy	788HG	20	QW576	56565655XX	0034



## APPENDIX D

*Report on access control*Figure 48: *Grid user administrative task options*

```

Your choice_ 3
CENTRAL SERVICES AND MONITORING UNIT
-----
WELCOME TO GRID COMPUTING -
(Resource Management)
Login as:
  1. Admin
  2. User
Q - QUIT

Your choice_ 1
Administrative task

1 - View all Domains and their roles
2 - View all Domains Users
3 - View all Roles and services
4 - View all info
5 - Back to main menu
Q - QUIT

Your choice_
-----
```



Figure 49: Available roles and services

```

...
Available Accessible Services
| Role                | Service                |
-----|-----|
| Cardiologist        | read patient record   |
| Cardiologist        | read prescription     |
| Cardiologist        | treat heart diseas   |
| Cardiologist        | write patient record  |
| Cardiologist        | write prescription    |
| Dentist             | carry out operation   |
| Dentist             | read patient record   |
| Dentist             | read prescription     |
| Dentist             | treat teeth           |
| Nurse              | read patient record   |
| Nurse              | read prescription     |
| Nurse              | write patient record  |
| Ultrasound Technologist | analyse images       |
| Ultrasound Technologist | read patient record   |
| Ultrasound Technologist | take ultrasound       |
-----|-----|
Administrative task

1 - View all Domains and their roles
2 - View all Domains Users

```



Figure 50: Grid users' roles, domain hierarchy and their access status.

```

ii
  44i      Your choice_ 4
iisii_ displayAllUserQuery() ...
iisii
iisii| DomainID | DomainHeirachy | Name      | Surname | Role                | Remarks | Status |
iisii=====
iisii
iisii| A      | 1      | Saheed   | Dames   | Cardiologist        | invalid | terminate |
iisii| B      | 2      | James   | Kolade  | Dentist              | valid  | proceed  |
iisii| C      | 3      | Felix   | Adams   | Ultrasound Technologist | invalid | terminate |
iisii| B      | 2      | Daves   | Jill    | Nurse                | valid  | proceed  |
iisii| B      | 2      | Kane    | Peter   | Dentist              | invalid | terminate |
iisii| A      | 1      | Salack  | Jeleel  | Cardiologist        | valid  | proceed  |
iisii| A      | 1      | Agbele  | Kehinde | Pathologist          | valid  | proceed  |
iisii| A      | 1      | Daniel  | Lindert | Pulmonologist       | invalid | terminate |
iisii| B      | 2      | Kola    | Serifa  | Surgeon              | valid  | proceed  |
iisii| C      | 3      | Malixa  | Jenifa  | Obstetrician        | invalid | terminate |
iisii| B      | 2      | Andy    | Lui     | Oncologist          | valid  | proceed  |
iisii| B      | 2      | Apalara | Tijani  | Physician            | invalid | terminate |
iisii| A      | 1      | Sheick  | Caled   | Dental assistant    | invalid | terminate |
iisii| A      | 1      | Nurayhn | Wale    | Podiatrist          | valid  | proceed  |
  
```

The diagram shows two ovals on the right side of the table. The top oval is labeled "Permission denied" and has arrows pointing to it from the "Remarks" column of rows 1, 3, 4, and 5. The bottom oval is labeled "Permission granted" and has arrows pointing to it from the "Status" column of rows 2, 6, 7, and 8.

Figure 51: Information about the users, domain and role

DomainID	UserID	Name	Surname	Role	Remarks	Status
A	1	Frederick	Ishimwe	Cardiologist	valid	proceed
B	2	Azeez	Nureni	Dentist	valid	proceed
C	3	John	Doe	Ultrasound Technologist	valid	proceed
B	4	Jack	Jill	Nurse	valid	proceed
B	5	Mathew	Peter	Dentist	valid	proceed
A	6	Jackson	Jonson	Cardiologist	valid	proceed
C	7	Bello	Lukman	Dentist	valid	proceed
A	8	Azeez	Ahmed	Nurse	valid	proceed
C	9	Julius	Cliford	Ultrasound Technologist	valid	proceed
A	10	Ademola	Abidoeye	Cardiologist	valid	proceed
B	11	Isabella	Venter	Nurse	valid	proceed

Administrative task

- 1 - View all Domains and their roles
- 2 - View all Domains Users
- 3 - View all Roles and services
- 4 - View all info
- 5 - Back to main menu
- Q - QUIT

## APPENDIX E

*Sample schemas for describing interoperability using a template presented in Figure 16.*

The following sample schemas are used. The schemas contain sharable and relevant information from the three domains as shown in Figure 16.

<b>Register</b>	StdNo	SName	DOB				
<b>Student</b>	StdNo	Address	Gender				
<b>PostGradStd</b>	StdNo	SName	FName	Major	CGPA	Advisor	Workload
<b>UGStudent</b>	StdNo	SName	FName	GPA			
<b>StdProjec</b>	ProjNo	Supervisor	Report				
<b>Graduation</b>	GradDate	StdNo	GradVenue				

*Figure 52: Sample schema for Domain A (University)*

In schema shown in Figure 52, relation Student has StdNo as the primary key and attributes Address and Gender to describe the address of the student as well as his gender.

Relation Register has the Surname (SName), First Name (FName) and Date of Birth (DOB) as information for student (previously identified by StdNo) is registered.

PostGradStd, Student, StdProjec t and Graduation contain specific information about student.



<b>Register</b>	PatientID	SName	FName	
<b>Patient</b>	PatientID	Sex	Address	
<b>PatNeed</b>	Emergency	PatientID	Major	telephone
<b>PatDescription</b>	MedicalAid	HealthCondition	Payment	

Figure 53: Sample schema for Domain B (Hospital)

Also, in the sample schema presented in Figure 53, for Domain B (Hospital) relation Register has a “PatientID” as the primary key while SName, FName and DoB are attributes. Other relations include Patient, PatientNeed and PatDescription.

Sample schema for Domain C (Banking) is presented in Figure 54. Relation account has account-number as the primary key to attributes branchName and Balance. Branch, Customer and account-Holder are other relations with different attributes.

<b>Account</b>	branchName	Balance	AccountNumber		
<b>Branch</b>	branchName	Address	assets		
<b>Customer</b>	CustomerNumber	Name	Address	HomeBranch	
<b>Account-Holder</b>	CustomerNumber	AccountNumber	DoB		
<b>Transaction</b>	Transaction-ID	Account-ID	Currency	Amount	Balance

Figure 54: Sample schema for Domain C (Banking)

At this juncture, appropriate meta-queries were issued on schemas designed for Domains A, B and C. To achieve interoperability across these schemas, various operations such as EXTRACT, UPDATE, MERGE, DELETE and SELECT can be carried out on them.

In extracting information across the three domains, a meta-query which is of the form :

SELECT

From object Varnam, object Varname....

WHERE *page*

is considered.

Hence,

$X.attr - X[A].attr - Y[B].attr - Z[C]$

Where A, B and C are domains and attr-X, attr-Y and attr-z are attributes in each of the domains.

What has been achieved by the above statement is that all the attributes specified for all the domains A, B and C have been aggregated as unitary information. This implies that the aggregated information can now be used by any grid user.

Also,

SELECT \*  
FROM  
WHERE Register.StdNo [A] .Patient[B].attr-Y. ACCOUNT-HOLDER [C]

Here, the procedure “Extract” will bring all the information in relations “Register”, “Patient” and “account-holder” that belong to domains A, B and C.

In order to merge different schemas from the three domains, the following meta-query language was issued.

**SELECT** PostGradstd.StdNo , Graduaton.GradDate FROM Domain A

**UNION**

**SELECT** Register.PatientID , PatDescription.MedicalAid FROM Domain B

**UNION**

**SELECT** Account.AccountNumber, Account-Holder.CustomerNumber FROM  
Domain C

From the foregoing, it has been established how interoperability is achievable when schemas are created across Domains.

## APPENDIX F

*The means and the standard deviations of the experiments shown in Figures 29 and 30.*

Definitions:

1. Let  $M_{SB}$  denotes variance between the three domains considered
2. Let  $M_{SW}$  denotes variance within the three domains considered

To evaluate both the means and standard deviation of the experiment shown in Figure 29, we construct hypothesis test based on the values obtained using ANOVA.

$H_0: \mu = \mu_A = \mu_B = \mu_C$ , where A, B and C are domains considered.

$H_1$ : At least one of the mean is different from the others.

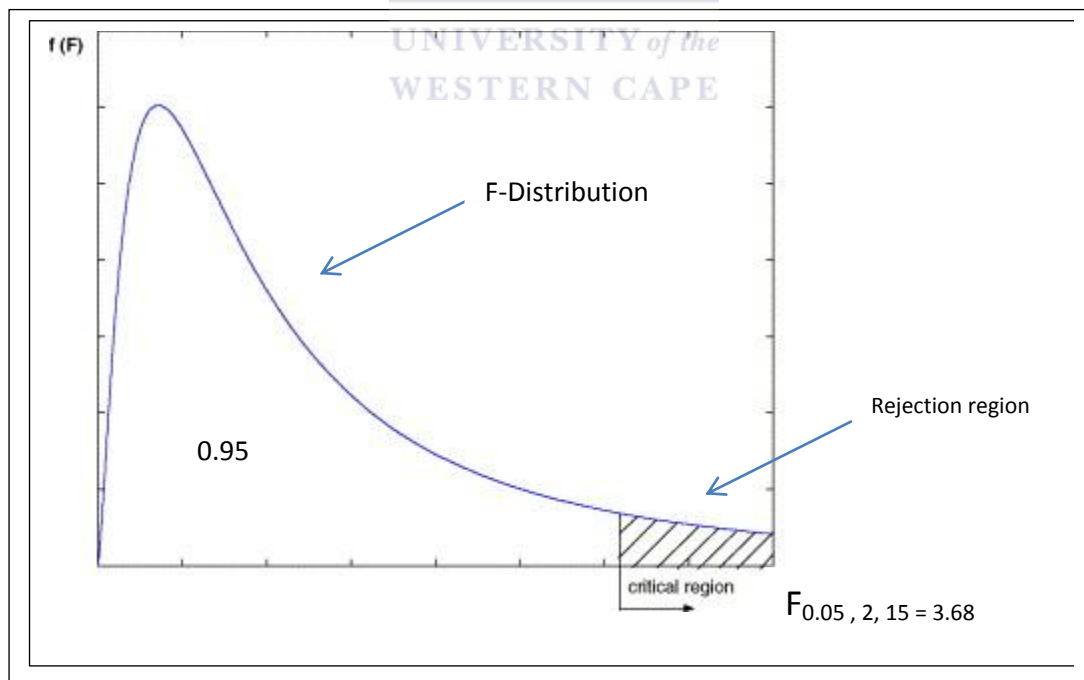


Figure 55: Showing values of 3.68 at  $F_{0.05, 2, 15}$

It is noted that there are currently the value of  $K = 3$  domains, that is, Domains A, B, C. Therefore,  $DoF_N = K - 1 = 3 - 1 = 2$ . The sum of data for all the three domains denoted as  $N = n_1 + n_2 + n_3 = 6 + 6 + 6 = 18$ .

Using the  $DoF_D = N - K = 18 - 3 = 15$  and  $\alpha = 0.05$  (the Least Significant value).

The critical value if  $F_{0.05, 2, 15} = 3.68$  (determined using F-Distribution table)

There is need to find :  $\bar{\bar{x}} = \text{mean of mean} = \sum X/N$ ,  $M_{SB} = \sum ni (\bar{X} - \bar{\bar{x}})^2 / K - 1$  and  $M_{SW} = \sum (ni - 1) S_i^2 / K - 1$

Table 20: The values for average Turnaround time and the No of grid nodes for Domains A, B, C

Parameters Determined	No of grid nodes	Turnaround time (T) (Domains A,B, C)			Total Turnaround time for Domains A,B, C	Average Turnaround time for Domains A,B, C
		Domain A	Domain B	Domain C		
	10	12 = 30.77%	14 = 35.89%	13=33.33%	39	13
	20	7.0 =29.17%	8.0 =33.33%	9.0=37.50%	24	8
	30	4.2=23.33%	6.4=35.55%	7.4=41.11%	18	6
	40	3.8=25.33%	4.8=32.00%	6.4=42.67%	15	5
	50	3.1=25.83%	3.9=32.50%	5.0=41.67%	12	4
	60	1.4=23.33%	2.0 =33.33%	2.6=43.33%	6.0	2
$\sum X$	→	31.5	39.1	43.4		
$\bar{X}$	→	5.25	6.51	7.233		
$s^2$	→	11.879	14.73	10.59		
n	→	6	6	6		
						N = 18

The mean of mean denoted as  $\bar{\bar{x}}$  was determined as follows:

$$\bar{\bar{x}} = 31.50 + 39.10 + 43.40 / 18 = 6.33,$$

The mean for each of the domains are evaluated as follows:

$$\bar{X}_{\text{Domain A}} = \sum X / n, = 31.5 / 6 = 5.25, \bar{X}_{\text{Domain B}} = \sum X / n = 39.1 / 6 = 6.51 \text{ and } \bar{X}_{\text{Domain C}} = \sum X / n = 43.40 / 6 = 7.233$$

$$s^2_{\text{Domain A}} = 1/N \sum X_i^2 - \bar{X}^2 = 236.65 / 6 - 5.25^2 = 11.879, \text{ while } s^2_{\text{Domain B}} = 14.73 \text{ and } s^2_{\text{Domain C}} = 10.59$$

$$\text{Mean of Mean ; } \bar{\bar{x}} = (31.5 + 39.1 + 43.4) / 18 = 6.33$$

Also, from Table 20 shown,  $M_{SB} = \sum ni (\bar{X} - \bar{\bar{x}})^2 / K-1$  could be determined as follows:

$$= 6(5.25 - 6.33)^2 + 6(6.51 - 6.33)^2 + 6(7.233 - 6.33)^2 / 2$$

$$= 6.042$$

Also,

$$M_{SW} = \sum (ni - 1) S_i^2 / N-K$$

$$= (6-1) 11.879 + (6-1) 14.73 + (6-1) 10.59 / 15$$

$$= 12.399$$

Therefore, the test statistics is  $F = M_{SB} / M_{SW} = 6.042 / 12.399 = 0.4872$

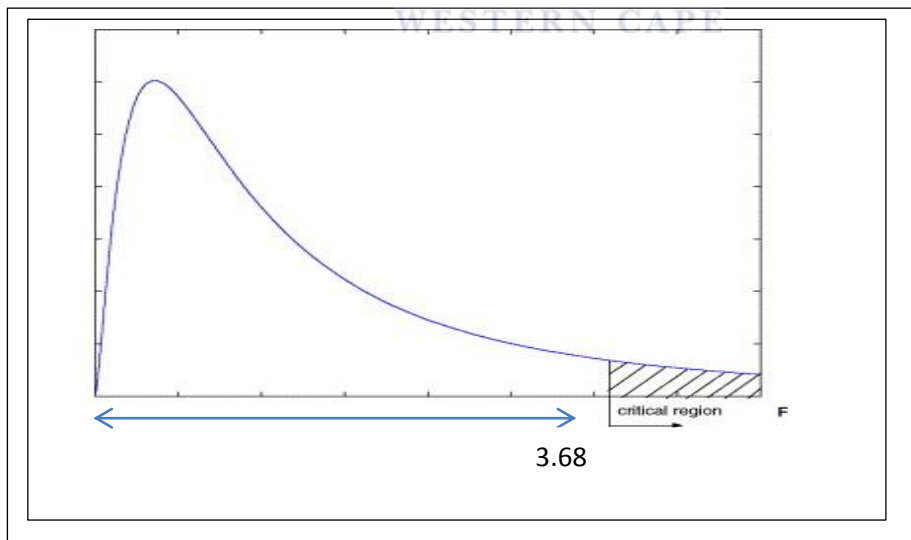


Figure 56: Showing F-Distribution table for 3.68

**Conclusion:**

Since F-Statistical table falls to the left of F-distribution ( $0.4872 < 3.68$ ) therefore, we may conclude that there is no significance difference in the means of group of Domains A,B, and C. There is significantly no difference in the means even though the values of the grid nodes were varied independently across each of the domains and the corresponding values of the turnaround time.

The graph of the No of grid nodes with their corresponding average turnaround time is presented in Figure 29 based on the values analysed in Table 20.

Table 21: The values for average Turnaround time and the No of grid service requesters for Domains A, B, C

Parameters Determined	No of service requester	Turnaround time (T) (Domains A,B, C)			Total Turnaround time for Domains A,B, C	Average Turnaround time for Domains A,B, C
		Domain A	Domain B	Domain C		
	50	1.1 = 36.66%	1.4 = 46.66%	0.5 = 16.66%	3.0	1.0
	100	2.6 = 30.95%	3.1 = 36.90%	2.7 = 32.14%	8.4	2.8
	200	2.8 = 29.17%	3.4 = 35.41%	3.4 = 35.42%	9.6	3.2
	300	3.0 = 29.41%	3.6 = 35.29%	3.6 = 35.29%	10.2	3.4
	400	3.6 = 30%	4.4 = 36.67%	4.0 = 33.33%	12.0	4.0
$\sum X$	→	13.1	15.9	14.2		
$\bar{X}$	→	2.62	3.18	2.84		
$s^2$	→	0.6896	0.9776	1.5464		
n	→	5	5	5		

N = 15

To find both the mean ( $\bar{X}$ ) and the standard deviation ( $s^2$ ) of experiment whose result is graphically presented in Figure 30, hypothesis was conducted based on the data obtained during experimentation.

As usual, the values of  $K$ ,  $DoF_N$ ,  $N$  and  $DoF_D$  are determined as :  $K = 3$ ,  $DoF_N = K - 1 = 3 - 1 = 2$ ,  $N = n_1 + n_2 + n_3 = 5 + 5 + 5 = 15$  while  $DoF_D = N - K = 15 - 3 = 12$  and  $\alpha = 0.05$  (the Least Significant value).

For the experiment whose values appear in Table 21, the critical value  $F_{0.05, 2, 3.89}$  (using F-Distribution table) is determined.

$$s^2_{\text{Domain A}} = 1/N \sum X_i^2 - \bar{X}^2 = 37.77/5 - (2.62)^2 = 0.68644, \quad s^2_{\text{Domain B}} = 0.9776 \quad \text{and} \quad s^2_{\text{Domain C}} = 1.5464$$

To calculate the Mean of Mean ( $\bar{\bar{x}}$ ) based on the values provided in the Table 21, therefore,

$$\bar{\bar{x}} = 13.1 + 15.9 + 14.2 / 15 = 2.88$$

Also, from Table 21 shown,  $M_{SB} = \sum ni (\bar{X} - \bar{\bar{x}})^2 / K - 1$  could be determined as follows:

$$\begin{aligned} &= 5 (2.62 - 2.88)^2 + 5(3.18 - 2.88)^2 + 5(2.84 - 2.88)^2 / 2 \\ &= 0.796 / 2 = 0.398 \end{aligned}$$

Also,

$$\begin{aligned} M_{sw} &= \sum (ni - 1) S_i^2 / K - 1 \quad \text{was obtained as follows:} \\ &= (5-1) 0.6896 + (5-1) 0.9776 + (5-1) 1.5464 / 12 \\ &= 12.8544 / 12 \\ &= 1.0712 \end{aligned}$$

Therefore, the statistics  $F = M_{SB} / M_{sw}$

$$\begin{aligned} &= 0.398 / 1.0712 \\ &= 0.3715 \end{aligned}$$

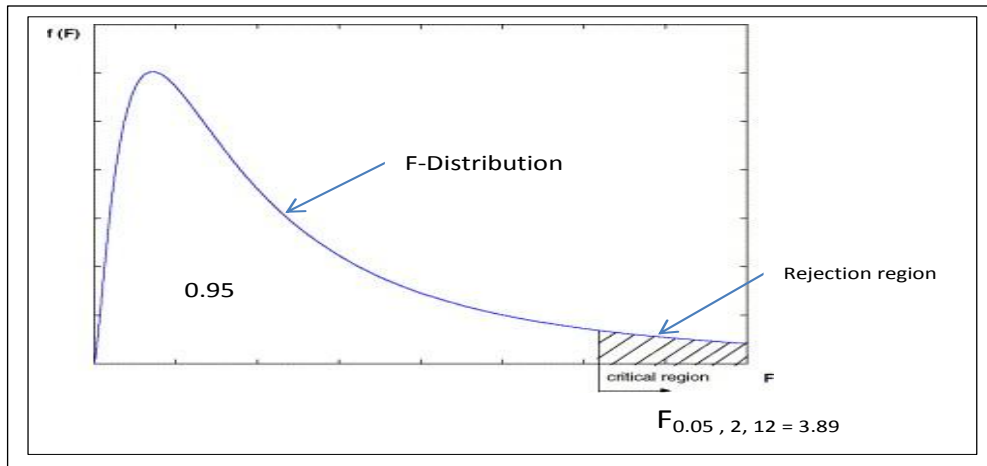


Figure 57: Showing values of 3.89 at  $F_{0.05, 2, 12}$

Finally, since F-Statistical table falls to the left of F-Distribution ( $0.3715 < 3.89$ ), it can therefore be concluded that there is no significance difference in the Mean of group of Domains A, B, C. Invariably, there is significantly no difference in the Mean even though; the values of No of grid service requesters were varied independently across each of the domains and the corresponding values of the average turnaround time. The final graph is presented in Figure 30 where the average turnaround time is plotted against the number of grid service requesters.



## APPENDIX G

*Sample source code for implementing access control*

```

import java.io.*;
import java.sql.*;

public class testResourceMngt
{
    static BufferedReader inKb = new BufferedReader (new
InputStreamReader (System.in));
    static Gridatabase db = new Gridatabase();
    static char choice = ' ';
    static boolean loggedin =false;

    public static void login() throws IOException,SQLException{
        String uname = "",pass = "";
        System.out.println("\n\n      Grid User Login\n");
        System.out.println("      =====\n");
        System.out.print("      Username:");
        uname =inKb.readLine();
        System.out.print("\n Password:");
        pass =inKb.readLine();
        System.out.println();
        db.isUser (uname,pass);

        loggedin =true;
    }
    public static void admin() throws IOException,SQLException
    {
        do
        {
            System.out.println("      Administrative task");
            System.out.println();
            System.out.println("      1 - View all Domains and their roles");

```

```
System.out.println("    2 - View all Domains Users ");
System.out.println("    3 - View all Roles and services ");
System.out.println("    4 - View all info ");
System.out.println("    5 - Back to main menu ");

System.out.println("    Q - QUIT");
System.out.println(" ");
System.out.print("    Your choice_ ");
choice =inKb.readLine().toUpperCase().charAt(0);

switch(choice)
{
    case '1':
    {
        db.view_Domain_Roles();
        break;
    }
    case '2':
    {
        //Print all Domains Users
        db.view_Domain_user();
        break;
    }
    case '3':
    {
        //Print all Domains Users
        db.view_role_service();
        break;
    }
    case '4':
    {
        //Print all Domains Users
        db.displayAllUserQuery();
        break;
    }
    case '5':
    {
        //Print all Domains Users
        menu();
        break;
    }
}

}while (choice!='Q');
}
public static void user() throws IOException,SQLException
{

System.out.println("Current User details");
```

```

        System.out.printf("USER \t: %s\nDOMAIN\t: %s\nROLE \t:
%s\n",db.getName(),db.getDomain(),db.getRole());

do
{
    System.out.println("    Access Resources and servives");
    System.out.println("=====");

    System.out.println("    1 - Show Accessible
Resources/Services");
    System.out.println("    2 - Show all services");
    System.out.println("    3 - Back to main menu ");

    System.out.println("    Q - QUIT");
    System.out.println(" ");
    System.out.print("    Your choice_ ");
    choice = inKb.readLine().toUpperCase().charAt(0);

    switch(choice)
    {
        case '1':
        {
            //Show Accessible Resources/Services
            db.allowedService();
            break;
        }
        case '2':
        {
            //Show all services
            allserviceQuery();

            break;
        }
        case '3':
        {
            //back to main menu
            menu();
            break;
        }
    }

    }while (choice!='Q');
inKb.close();
}

public static void allserviceQuery() throws SQLException,IOException
{
    String option = "", access = "";

```

```

do
{
    db.allService();
    System.out.println("\n      Select service :");
    System.out.println("      Q - QUIT");
    System.out.println(" ");
    System.out.print("      Your choice_ ");
    option = inKb.readLine();
    choice =option.toUpperCase().charAt(0);
    access =db.verifyService(option);
    if(!access.equals("Granted")){
        System.out.printf("Access %s !!\n You are not Authorised to
use this service in your domain (%s)\n",access,db.getDomain());}
        else{System.out.printf("Access %s !!\nProceed ... \n",access);}
    }
while (choice != 'Q' );
    inKb.close();
    // DB.disconnect();
    System.out.println("Done");
}

public static void menu()throws SQLException,IOException
{
    System.out.println();
    System.out.println("CENTRAL SERVICES AND MONITORING
UNIT\n_____");
    System.out.println();
    do
    {
        System.out.println("      WELCOME TO GRID COMPUTING -
\n(Resource Management) ");
        System.out.println("Login as: \n \t 1. Admin \n \t 2. User");
        System.out.println("      pQ - QUIT");
        System.out.println(" ");
        System.out.print("      Your choice_ ");
        choice = inKb.readLine().toUpperCase().charAt(0);
        switch(choice)
        {
            case '1':
                {
                    if(!loggedin)
                        login();//Admin
                    admin();
                    break;
                }
            case '2':
                {
                    if(!loggedin)
                        login(); // user
                    user();
                }
        }
    }
}

```

```
                break;
            }
        }
    }
    while (choice != 'Q' );

    inKb.close();
    // DB.disconnect();
    System.out.println("Done");
}

public static void main (String[] args) throws
SQLException, IOException
{
    // NewsPaper91 DB = new NewsPaper91();
    menu();
    db.disconnect();
} // main

} // class
```



PUBLICATIONS AND CONFERENCE PROCEEDING





UNIVERSITY *of the*  
WESTERN CAPE